

EVALUACIÓN HEURÍSTICA Y PRUEBAS CON USUARIOS: APLICACIÓN A INTERFACES DE BASES DE DATOS TERMINOLÓGICAS

Mesa Lao, Bartolomé. Profesor asociado. Departamento de Traducción e Interpretación. Facultad de Traducción e Interpretación. Universidad Autónoma de Barcelona, Edificio K, 08193 Bellaterra (Barcelona), barto.mesa@uab.es; Marcos, Mari-Carmen. Profesora lectora. Instituto Universitario de Lingüística Aplicada (IULA), Universidad Pompeu Fabra (Barcelona), La Rambla 30-32, 08002 Barcelona, mcarmen.marcos@upf.edu; Morales Moreno, Albert. Becario IULA. Instituto Universitario de Lingüística Aplicada (IULA), Universidad Pompeu Fabra (Barcelona), La Rambla 30-32, 08002 Barcelona, albert.morales@upf.edu; Mayor Lloret, Agustí. Máster de Lingüística Aplicada. Instituto Universitario de Lingüística Aplicada (IULA), Universidad Pompeu Fabra (Barcelona), La Rambla 30-32, 08002 Barcelona, agusti.mayor.lloret@gmail.com.

Resumen: A menudo los usuarios se enfrentan a sistemas de recuperación de información que no resultan fáciles de entender ni de utilizar. Este es el caso de los profesionales de la lengua (traductores, correctores, etc.) que deben emplear a diario herramientas de consulta de información lingüística para poder realizar su trabajo.

Presentamos un estudio realizado por un grupo de investigadores de la universidad Pompeu Fabra entre diciembre de 2005 y marzo de 2006, en el que se ha desarrollado una metodología de análisis de calidad de bases de datos terminológicas tomando como criterio principal la usabilidad. Los resultados obtenidos se basan en una metodología de evaluación heurística, complementada posteriormente con pruebas realizadas con usuarios: tests, cuestionarios y entrevistas.

Nuestra investigación toma como objeto de análisis diez bases de datos terminológicas cuya consulta es libre a través de Internet. Todas pertenecen a diferentes ámbitos temáticos y ofrecen opciones de consulta multilingüe. De los resultados obtenidos se deduce que las bases de datos terminológicas no están considerando pautas importantes de usabilidad.

Abstract: *In the day-to-day work with information retrieval tools, users are faced with systems that are difficult to understand and use. This is the case for linguists and translators, experts in language but not necessarily in the use of all these reference tools.*

In this article we present a study carried out between December of 2005 and March of 2006 by a group of researchers from the Pompeu Fabra University in which we established and applied a methodology to analyze the quality of ten terminology databases from the usability point of view.

Results are based on heuristic evaluation and user testing (tests, questionnaires and interviews) to show that terminology databases on the Internet are not fully considering usability guidelines (navigation/architectural clarity and effectiveness/efficiency/satisfaction).

Palabras clave: bases de datos terminológicas, evaluación centrada en el usuario, evaluación heurística, interacción persona-ordenador, interfaces de usuario, sistemas de recuperación de información, terminología, usabilidad.

Keywords: *heuristic evaluation, human-computer interaction, information retrieval systems, terminology, terminology databases, usability, user-centered evaluation, user interfaces.*

Introducción

En su quehacer diario, traductores, correctores y lingüistas manejan herramientas lexicográficas y terminológicas variadas, entre ellas las bases de datos terminológicas, que son el objeto de estudio de esta comunicación. Al igual que ocurre en otros ámbitos, el manejo de sistemas informatizados puede resultar un obstáculo en vez de una ayuda. A menudo la complejidad de estos sistemas impide a los usuarios centrarse en sus objetivos, haciendo que deban prestar atención al funcionamiento del sistema y mermando, por ende, la atención a las tareas propias de su profesión.

Este problema no resulta una novedad, sino que viene dándose desde los inicios de la informática, pues tradicionalmente se han diseñado sistemas tomando como punto central la tecnología disponible y no las capacidades ni las necesidades de sus usuarios finales, que son realmente la razón de ser de su existencia. Por este motivo, los usuarios precisan de una etapa de aprendizaje de los sistemas de recuperación de información para aprovechar al máximo las funcionalidades que ofrecen estos sistemas. La disciplina denominada interacción persona-ordenador (IPO) surge en los años 50 precisamente para estudiar los problemas de comunicación entre las máquinas y las personas, y determinar buenas prácticas que mejoren esta interacción. Esta disciplina toma como fuente de conocimiento la psicología cognitiva, la ergonomía y la ingeniería informática. A pesar de los grandes avances logrados y de estar extendiéndose el concepto de usabilidad y el de arquitectura de la información, todavía queda mucho camino por recorrer tanto en aplicaciones de escritorio como en la web, por no contar los dispositivos ubicuos (teléfonos móviles, PDA, etc.).

En el caso de las bases de datos terminológicas disponibles en Internet la situación no es diferente. Un estudio realizado por Marcos et al. (2006) sobre diez sistemas online de acceso gratuito ponía de manifiesto las carencias de estas herramientas para poder ser consideradas «usables», es decir, fáciles de comprender y de utilizar por los usuarios para los que han sido diseñadas. Para realizar este análisis se partió de las bases heurísticas establecidas por Nielsen (1994, 2000, 2002a, 2002b) y Nielsen & Molich (1990) de forma general, por Marcos & Cañada (2003), Marcos (2004) para sistemas de recuperación de información, y por Marcos & Gómez (2006) de forma más específica para sistemas terminológicos. A partir de los resultados del análisis heurístico y de las pruebas realizadas con usuarios se llegó a la conclusión que la navegación por estos sistemas puede optimizarse.

Cuando se trata de evaluar un producto o sistema de información, resulta obvio decir que el contenido es un punto clave para medir su calidad, pero no es el único, pues los aspectos de forma también son importantes. En este trabajo analizamos problemas de navegación y claridad arquitectónica (disposición sistemática y coherente de los contenidos) detectados en las bases de datos terminológicas seleccionadas con la intención de poner de manifiesto la necesidad de diseñar los sistemas teniendo en cuenta las necesidades de los usuarios que las utilizaran. Para ello contamos con información obtenida de análisis heurísticos y de pruebas con usuarios complementadas con los cuestionarios y entrevistas que se les realizaron.

Objetivos

Partiendo del análisis de diez bases de datos terminológicas multilingües en Internet, se plantean los siguientes objetivos:

1. Examinar a través de una metodología de análisis heurístico el grado de usabilidad de diez bases de datos terminológicas desde el punto de vista de la navegación y la claridad arquitectónica. En este sentido se evalúa el grado de usabilidad en relación con la facilidad de movimiento por las diferentes bases de datos a la hora de realizar consultas y recuperar resultados.
2. Analizar el grado de eficiencia, eficacia y satisfacción que presentan estos sistemas a través de una metodología de análisis basada en pruebas con usuarios (tests, entrevistas y cuestionarios). Para el análisis se han seleccionado estos tres aspectos, ya que son los tres aspectos que recoge la definición de usabilidad de la ISO (1).

Metodología

1.1 Muestra de análisis

Las bases de datos sobre las que se ha trabajado son algunas de las más populares entre el perfil de usuarios habituales de estos sistemas (traductores, redactores técnicos y asesores lingüísticos). Todas permiten recuperar información terminológica en más de una lengua y pueden consultarse de manera gratuita a través de Internet:

1. *Base de Terminologie*. <<http://www.cilf.org/bt.fr.html>>
2. *CercaTerm*. <<http://www.termcat.net>>
3. *Eurodicautom*. <<http://europa.eu.int/eurodicautom>>
4. *EuskalTerm*. <<http://www1.euskadi.net/euskalterm>>
5. *Oncoterm*. <<http://www.ugr.es/~oncoterm/alpha-index.html>>
6. *TerminoBanque*. <<http://www.cfwb.be/franca/bd/bd.htm>>
7. *TIS: Terminological Information*. <<http://tis.consilium.eu.int>>
8. *UBTerm*. <<http://www.ub.edu/slc/ubterm>>
9. *UNTerm*. <<http://unterm.un.org>>
10. *WTOTerm*. <<http://wtoterm.wto.org>>

1.2 Evaluación heurística

La evaluación heurística de la usabilidad es el método de análisis según el cual un conjunto de expertos evalúan un sistema detenidamente y describen los problemas que detectan. A continuación presentamos las 16 variables que hemos considerado para evaluar principios genéricos de navegación y claridad arquitectónica en las bases de datos terminológicas anteriores:

1. ¿Presentan todas las páginas un título identificativo?
2. ¿El enlace a la página principal (*homepage*) se identifica claramente?
3. ¿Los apartados más importantes del sitio son accesibles desde la página principal?
4. ¿Evita el sitio el paso forzoso por una página de bienvenida siempre que se visita (una página previa a la que contiene la caja de consultas)?
5. ¿La tarea que se está llevando a cabo en cada momento se identifica claramente (en qué fase de la recuperación de información terminológica nos encontramos)?
6. ¿Aparece la imagen corporativa del sitio en todas las páginas?
7. ¿Existe la posibilidad de realizar búsquedas avanzadas sobre la base de datos (operadores booleanos, especificación de la/s lengua/s de consulta, etc.)?
8. ¿El sistema responde de algún modo ante búsquedas vacías o incompletas?
9. ¿Existe la posibilidad de ordenar la lista de resultados obtenidos (por frecuencia de uso, por orden alfabético, etc.)?
10. ¿Los resultados de una búsqueda, en caso de que sean numerosos, aparecen paginados?
11. Además de la caja de texto para interrogar la base de datos, ¿ofrece el sistema otros sistemas de recuperación de terminología (navegación por listados de términos ordenados alfabéticamente, por ejemplo)?
12. ¿El diseño y la disposición de los contenidos es concisa y clara (se ha evitado la sobrecarga informativa)?
13. ¿La organización de la página se consigue con títulos, listas y una estructura constante?
14. ¿Todos los enlaces pueden reconocerse fácilmente e indican claramente su estado?
15. ¿El usuario no debe realizar el esfuerzo de recordar los datos ya introducidos en el sistema?
16. ¿La caja de texto para introducir las consultas en la base de datos está siempre visible?

Con el fin de lograr un análisis heurístico completo y con parámetros uniformes para todos los evaluadores (en el estudio participaron 5 evaluadores expertos), cada una de estas variables se ha puntuado mediante una plantilla en la que los expertos asignaban a cada variable una de estas tres opciones: «sí, siempre», «no, nunca» o «a veces». Responder «sí, siempre» suponía presencia de usabilidad en el sistema de recuperación de información terminológica y responder «no, nunca» suponía ausencia de usabilidad.

1.3 Evaluación con usuarios

El estudio de usabilidad con usuarios pone a prueba los sistemas observando la forma en que los usuarios trabajan con ellos. Este tipo de análisis permite obtener información objetiva y subjetiva del usuario, y comprobar de una manera más fiable si la base de datos evaluada resulta usable.

Para la elaboración de este estudio se ha contado con la colaboración de 30 usuarios, cuyo perfil medio responde a las siguientes características:

- Edad: entre 20 y 25 años.
- Formación o área profesional: estudios universitarios de ámbito lingüístico (estudios de Traducción e Interpretación, básicamente).
- Conocimientos de Internet: conocimientos avanzados.
- Antigüedad de navegación: entre 5 y 10 años.
- Horas diarias de navegación: más de 4 horas de navegación.
- Uso de Internet: finalidades profesionales (recurso de apoyo o medio para trabajar).
- Conocimientos sobre bases de datos terminológicas: avanzado.

El objetivo de este estudio con usuarios es analizar la eficiencia, eficacia y satisfacción con que trabajan los usuarios habituales de estos sistemas. Para ello se han utilizado distintas técnicas: el test de usuario, el cuestionario y la entrevista.

Mediante el test de usuario los observadores han podido tomar nota de las dificultades encontradas por los usuarios. Para ello, se les pidió que realizaran cuatro tareas de este tipo:

1. Busque en qué otros ámbitos de especialidad (medicina, economía, etc.) se usa el término X y qué equivalentes emplean otras lenguas para hacer referencia a este término.
2. Realice una búsqueda en la base de datos terminológica acotando el área temática de uso del término que busca. Por ejemplo, búsqueda de equivalentes terminológicos para el término *virus* sólo dentro del ámbito de la medicina, sólo dentro del ámbito de la informática, etc.
3. Recupere en la base de datos información relativa a X (no se le facilita un término de búsqueda concreto).
4. Consulte en la base de datos términos que empiecen por el radical X (uso de operadores de truncamiento, * / ?).

1.3.1 Eficiencia

Por medio de la observación directa basada en pruebas con usuarios, los evaluadores expertos pudieron determinar si los usuarios conseguían completar las tareas encomendadas; en concreto, se consideró que los sistemas resultaban efectivos si se conseguía localizar la información terminológica solicitada en las tareas.

1.3.2 Eficacia

La eficacia se evaluó a través del siguiente cuestionario de respuestas cerradas («sí», «no» y «a veces»):

1. ¿Le ha costado mucho tiempo dar con la respuesta o la ha encontrado rápidamente?
2. ¿Le parece eficaz la búsqueda?
3. ¿Le ha parecido fácil de usar la página de inicio?
4. ¿Ha realizado la búsqueda con rapidez y agilidad?

1.3.3 Satisfacción

El grado de satisfacción del usuario se evaluó mediante una entrevista cuyas repuestas se habían categorizado como «sí», «no» y «a veces».

1. ¿Le ha parecido fácil de usar la base de datos?
2. ¿Le ha resultado fácil interpretar los iconos, los textos de los menús y la estructura de las páginas?
3. ¿Resulta fácil la búsqueda en este sitio web?

Resultados

2.1. Resultados de la evaluación heurística

Las preguntas que en el análisis heurístico evalúan aspectos relacionados con la navegación por las bases de datos son las comprendidas entre los números 1 y 11 (ambas inclusive).

Las variables que mejores resultados han obtenido en el cuestionario han sido las relativas a la ordenación, claridad y gestión de los resultados mostrados (preguntas 9, 10, y 11, respectivamente). Las variables de análisis que han obtenido una puntuación media en los resultados hacen referencia al sitio web en el que se alberga la base de datos (preguntas 1, 2 y 4) y a la posibilidad de realizar búsquedas complejas (pregunta 7). Las variables relacionadas con la navegación que han obtenido peores resultados son las que hacen referencia a la accesibilidad a los apartados más importantes de la base de datos (pregunta 3) y a la contextualización de la sección o fase de búsqueda en la que nos encontramos en cada momento (pregunta 5). Se puede observar que dos de las variables más importantes en relación con la navegación por la base de datos (si no las más importantes, junto con las búsquedas avanzadas) son las que obtienen peores resultados.

En términos generales, los resultados obtenidos para evaluar la navegación por las bases de datos indican que:

- las bases de datos con mejores resultados son *UNTerm*, *Terminobanque* y *Euskalterm*,
- mientras que la base de datos con peores resultados es *Oncoterm*.

Las preguntas que en el análisis heurístico evalúan la claridad arquitectónica son las comprendidas entre los números 12 y 16 (ambas inclusive). Las preguntas 12 y 13 hacen referencia a la claridad arquitectónica de las bases de datos en relación con la estructuración de la información, mientras que las preguntas 14, 15 y 16 (que son las que han obtenido mejores resultados en la evaluación heurística) se refieren a la facilidad para realizar búsquedas y recuperar información en la base de datos.

En términos generales, los promedios obtenidos para el análisis de la claridad arquitectónica indican que:

- las bases de datos con mejores resultados en este sentido son *Euskalterm* y *Oncoterm*,
- mientras que la base de datos con peores resultados es *UBTerm*.

A continuación se presentan los resultados globales obtenidos para cada una de las diez bases de datos analizadas según la metodología del análisis heurístico. Las bases de datos que presentan puntuaciones más altas son las más usables desde el punto de vista de la navegación y la claridad arquitectónica. El promedio de ambas variables se recoge sobre una escala de 10 puntos. Para obtener las diferentes puntuaciones se han puntuado las variables de análisis presentadas anteriormente (§1.2) como 2, 1 o 0, según el resultado de las diferentes variables evaluadas fuese «sí, siempre», «a veces» o «no, nunca», respectivamente.

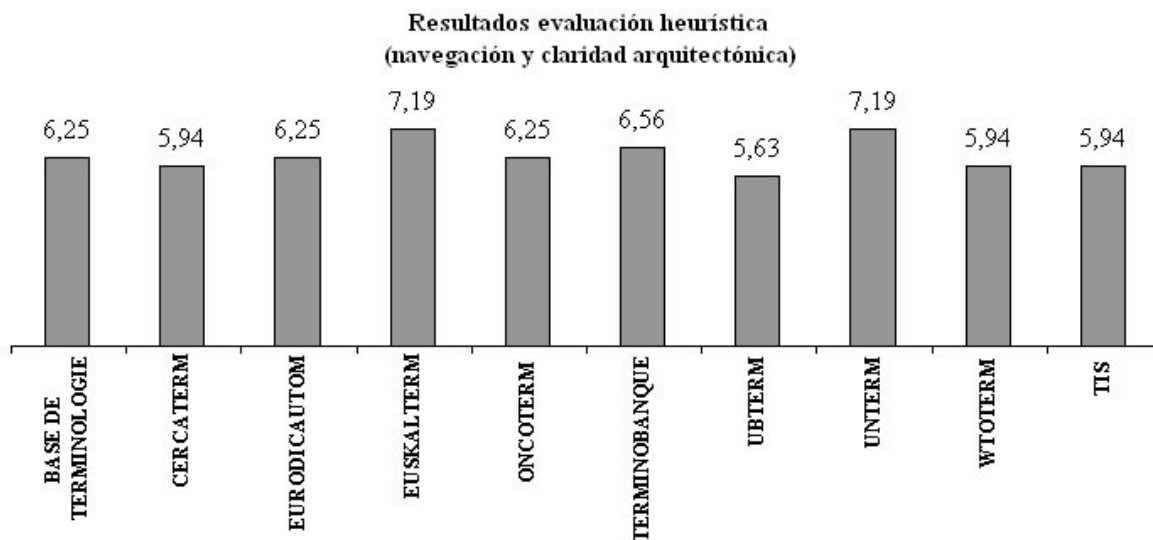


Figura 1. Resultados obtenidos a partir de la evaluación heurística para cada una de las bases de datos (escala de 10 puntos).

2.2. Resultados de la evaluación con usuarios

Los resultados obtenidos para la variable *eficiencia* son los siguientes:

- las bases de datos con mejores resultados son *Cercaterm*, *Eurodicautom* y *WTOTerm*,
- mientras que las bases de datos con peores resultados son *TIS* y *Terminobanque*.

Si analizamos las preguntas realizadas a los usuarios en relación con la eficiencia de las bases de datos, las preguntas 1 y 4 han sido las mejor valoradas. De estos resultados podemos deducir que las bases de datos cumplen el objetivo para el cual han sido diseñadas: localizar equivalentes terminológicos en diferentes lenguas. En el extremo opuesto vemos que el peor resultado fue obtenido en la tarea consistente en acotar el ámbito de especialidad en la búsqueda del término.

Los resultados obtenidos para la variable *eficacia* son los siguientes:

- las bases de datos con mejores resultados son *Base de terminologie*, *Oncoterm* y *UBTerm*,
- mientras que las bases de datos con peores resultados son *UNTerm*, *Terminobanque* y *Euskalterm*.

Si analizamos las preguntas realizadas a los usuarios para estas variables, las preguntas 1 y 4 han sido las mejor valoradas. Según estos resultados, deducimos que las bases de datos son eficaces desde el punto de vista de la ordenación interna, la claridad expositiva y la gestión de la información; lo que coincide con los resultados obtenidos en el análisis heurístico. Sin embargo, el peor resultado se ha obtenido en la segunda pregunta de la serie, la cual evalúa la eficacia de la base de datos.

Los resultados para la variable *satisfacción* son los siguientes:

- las bases de datos con mejores resultados son *TIS*, *Cercaterm* y *WTOTerm*,
- mientras que las bases de datos con peores resultados son *UNTerm* y *Eurodicautom*.

Los resultados muestran que los usuarios, en líneas generales, creen que las bases de datos son de fácil utilización; no obstante, han tenido dificultades con los iconos identificativos y con la estructura de las páginas.

Estos indicios corroboran los resultados del análisis heurístico de claridad arquitectónica y de navegación donde se advertían deficiencias en el acceso a los apartados más importantes de las bases de datos.

A continuación se presentan los resultados globales obtenidos para cada una de las diez bases de datos analizadas según la metodología de la evaluación con usuarios. Las bases de datos que presentan puntuaciones más altas son las más usables desde el punto de vista de la eficiencia, eficacia y satisfacción. El promedio de estas tres variables se recoge sobre una escala de 10 puntos.

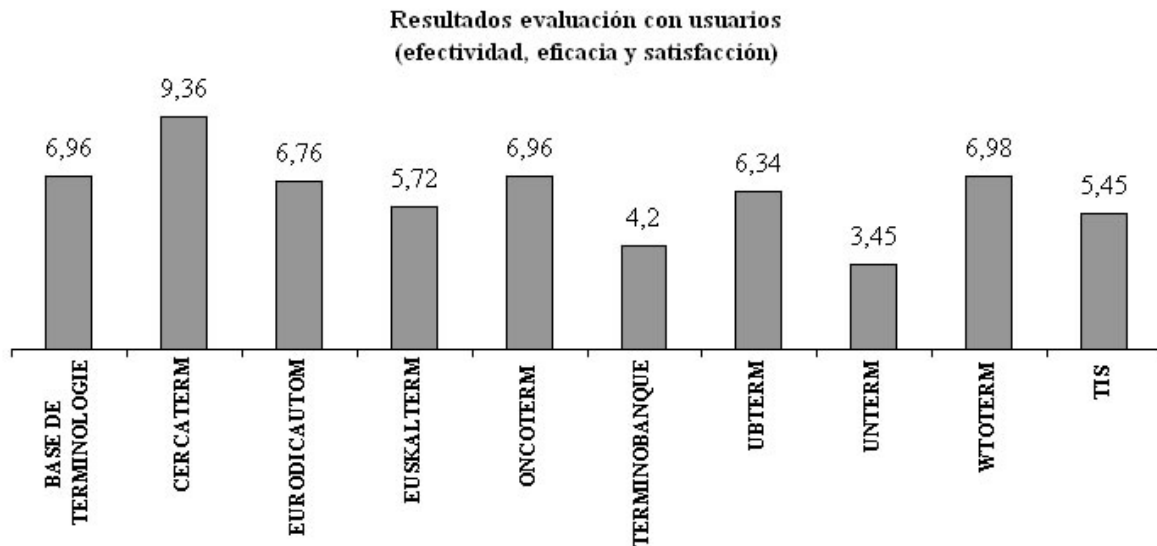


Figura 2. Resultados obtenidos a partir de la evaluación con usuarios (tests, cuestionarios y entrevistas) para cada una de las bases de datos (escala de 10 puntos).

Por último, como resultado de este estudio, se presenta un gráfico general en el que se pueden comparar sobre una escala de 10 puntos las evaluaciones realizadas por los expertos y por los usuarios para cada una de las bases de datos.

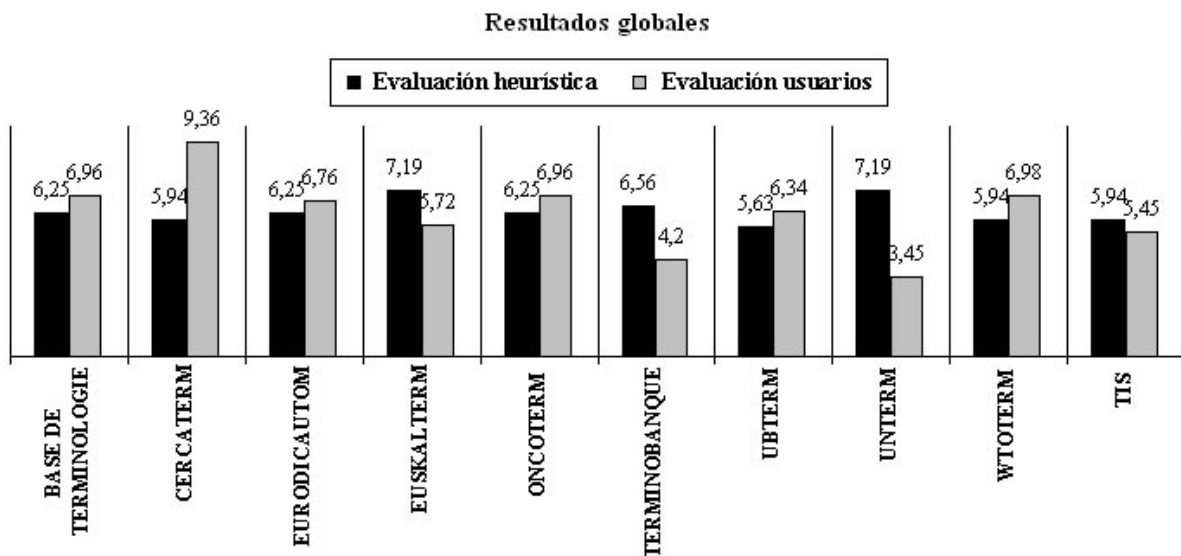


Figura 3. Comparativa de los resultados obtenidos mediante evaluación heurística y mediante pruebas con usuarios (escala de 10 puntos).

Conclusiones

Por un lado, el análisis heurístico permite concluir que:

1. Desde el punto de vista de la navegación, existen deficiencias en el diseño de las bases de datos relacionadas con aspectos de acceso y navegación por las secciones más importantes de las bases de datos.
2. Desde el punto de vista de la claridad arquitectónica, existen deficiencias en el diseño estructural del sitio web, ya que no se explota al máximo la pluriaccesibilidad a los contenidos. Los sistemas a menudo están orientados a representar la información de la base de datos de manera estática, sin facilitar opciones de personalización de consulta y acceso a dicha información según diferentes tipos de usuarios.

Entre todas las bases de datos analizadas, *Unterm*, *Euskalterm* y *Terminobanque* han sido las que han obtenido mejores puntuaciones generales en cuanto a navegabilidad y claridad arquitectónica.

Por el otro lado, el análisis con usuarios permite concluir que:

1. Desde el punto de vista de la eficiencia, las bases de datos terminológicas permiten realizar búsquedas con facilidad, tanto generales como avanzadas, desde cualquiera de las secciones de las bases de datos.
2. Desde el punto de vista de la eficacia, el análisis ha demostrado que, aunque las búsquedas sean rápidas y fáciles de realizar, los resultados no son los esperados por los usuarios. Sin embargo, no es nuestra intención evaluar aquí las bases de datos desde el punto de vista del contenido, sino de la forma.
3. Desde el punto de vista de la satisfacción, el análisis ha demostrado que el usuario experto no muestra satisfacción plena respecto a estos sistemas de recuperación terminológica y detecta a menudo deficiencias en el trabajo con las mismas. Esta situación creemos que se produce por la falta de unificación de criterios formales y de diseño en el diseño de bases de datos terminológicas.

Entre todas las bases de datos analizadas, *Cercaterm*, *WTOTerm* y *Base de terminologie* han sido las que han obtenido mejores puntuaciones generales en cuanto a eficiencia, eficacia y satisfacción.

Para finalizar diremos que el diseño de una aplicación interactiva —y, en concreto, de una base de datos terminológica— no es una suma de pantallas y de contenidos, sino que debe conceptualizarse como una herramienta de trabajo optimizada para el trabajo específico de un determinado perfil de usuarios. Por lo tanto, resulta imprescindible que los desarrolladores de este tipo de aplicaciones basadas en la recuperación de información terminológica sean conscientes de las características que definen la interacción persona-ordenador (IPO) en este campo, así como de los factores que determinan mejor la representación óptima de este tipo de información.

Técnicamente, el diseño de bases de datos usables favorece la seguridad, estabilidad, robustez y fiabilidad del sistema. Desde el punto de vista práctico, la usabilidad permite mayor rapidez en la realización de tareas y reduce las pérdidas de tiempo por parte del usuario. A partir de este breve estudio, corroboramos la necesidad de seguir implementando mejoras en la usabilidad de estas bases de datos terminológicas, lo que se traducirá por parte de los usuarios en:

- Mayor rapidez en la navegación.
- Mayor satisfacción y recompensa de uso.
- Mayor eficiencia en la difusión de la información que contienen.

Notas

- (1) «Usabilidad es la eficiencia, eficacia y satisfacción con la que un producto permite alcanzar objetivos específicos a usuarios específicos en un contexto de uso específico» [ISO/IEC 9241]

Referencias bibliográficas

- Marcos, M. C.** *Interacción en interfaces de recuperación de información: conceptos, metáforas y visualización*. Gijón: Trea, 2004
- Marcos, M. C.; Gómez, M.** “Idoneidad de las interfaces de léxicos y terminologías en la web”. En: *Glat: Aspects méthodologiques pour l'élaboration de lexiques unilingues et de multilingues*, 2006.
- Marcos, M. C.; Cañada, J.** “Cómo medir la usabilidad: técnicas y métodos para evaluar el uso de sitios web”. En C. Rovira y Ll. Codina (Dir.). *Documentación digital*. Barcelona: Sección Científica de Ciencias de la Documentación. Departamento de Ciencias Políticas y Sociales. Universidad Pompeu Fabra, 2003.
- Nielsen, J.** (1994). “Heuristic evaluation”. En J. Nielsen & R. Mack (eds.). *Usability Inspection Methods*. John Wiley & Sons, New York, NY.
- Nielsen, J.** How to Conduct a Heuristic Evaluation, 2002. Consultado en: 04-04-06.
http://www.useit.com/papers/heuristic/heuristic_evaluation.html
- Nielsen, J.** Ten Usability Heuristics, 2002. Consultado en: 04-04-06.
http://www.useit.com/papers/heuristic/heuristic_list.html
- Nielsen, J.** *Usabilidad. Diseño de sitios web*. Madrid: Prentice Hall, 2000.
- Nielsen, J. y Molich, R.** “Heuristic evaluation of user interfaces”, *Proceedings of the ACM CHI'90 Conference*, 1990, pp. 249-256.