

Las «listas negras» de palabras: un recurso informático poco conocido que facilita la coherencia terminológica*

M. Gonzalo Claros**

Resumen: La coherencia terminológica es uno de los aspectos en los que el traductor y el redactor cuentan con menos apoyo informático. En este artículo se va más allá de las correcciones ortográficas y se describe cómo se pueden utilizar las listas negras de términos para ayudarse a utilizar la terminología con coherencia. Cuando se trabaja con memorias de traducción, es posible hacer estas listas negras con Wordfast, pero las limitaciones más destacables de este pasan por no sugerir el término correcto, que solo se pueda usar una lista con cada documento y que se necesite un texto bilingüe. Sin embargo, estas limitaciones se pueden eliminar con la herramienta de listas negras de OpenOffice, que permite definir varias listas negras y usarlas a voluntad con cualquier documento, incluso los creados con la última versión de MS Word, sin ningún tipo de conversión ni de documento intermedio. Finalmente se sugiere un esquema de trabajo factible con control terminológico apto para cualquier traductor, revisor o redactor.

Palabras clave: lista negra, OpenOffice, coherencia terminológica, traducción asistida por ordenador, TAO.

“Blacklists” of words: a little known IT resource that facilitates terminology consistency

Abstract: Terminology consistency may be the issue with less IT support for translators and editors. This article goes beyond spelling corrections and describes how to use blacklists of terms to facilitate terminology consistency. Translators using translation memories can create these blacklists with WordFast, however, this program has some limitations, mainly because it does not suggest the correct term, it only allows one blacklist per document and it requires a bilingual text. The Blacklist Tool of OpenOffice surpasses these limitations by allowing the user to setup various blacklists and use them at will with any document, including those created with the latest version of MS-Word, without previous conversion or intermediate documents. A framework with feasible terminology control suitable for any translator, reviewer or editor is also provided.

Key words: blacklist, OpenOffice, terminology consistency, computer-assisted translation, CAT.

Panace@ 2009; 10 (29): 9-12

1. Introducción

A todos nos gustaría poder controlar la terminología de nuestros textos (redactados en español o traducidos a este idioma). Querríamos disponer de una herramienta que nos ayudase a mantener una coherencia interna, o bien que nos permitiera respetar los usos lingüísticos de un cliente. Quienes tenemos la necesidad de escribir ciencia en español también echamos de menos una herramienta que nos indique los términos o expresiones inaceptables que hemos trasladado a nuestro texto. Es cierto que, a menudo, se suele pedir a otra persona que nos revise el texto con ese objetivo, pero no siempre es posible o los plazos no siempre nos lo permiten. En cualquier caso, lo habitual es que quien escribe el texto y quien lo revisa crean que no tienen más ayuda que sus conocimientos y destrezas para la revisión terminológica. Esto enlaza con el hecho de que la mayoría de los redactores, traductores y revisores se conforman con sacar el mayor rendimiento posible a los ordenadores en relación con las correcciones ortográficas, pero no imaginan que también les pueden ayudar con la coherencia terminológica.

De todos es sabido que los distintos procesadores de texto del mercado (el omnipresente MS Word, pero también Open-

Office, WordPerfect, AbiWord, Pages, Mellel, TextEdit, etc.) ofrecen un corrector ortográfico en distintos idiomas y un diccionario personal donde incluir todas aquellas palabras que no vienen en los diccionarios suministrados por el fabricante. Puesto que la mayoría de los clientes exigen que los textos definitivos no tengan errores detectables por dichas herramientas, se presupone que el lector de este artículo no tiene problemas con su manejo. Tampoco entraremos a describir las gracias y desgracias de los diccionarios bilingües que vienen con los programas de ayuda a la traducción, puesto que solo son útiles cuando se usan, pero no, por ejemplo, cuando se escribe un texto directamente en español. En consecuencia, en este artículo trataremos situaciones como las siguientes.

- 1) Cada cliente tiene sus preferencias de uso para términos correctos. Por ejemplo, uno puede preferir *periodo* a *periodo*, *cardiaco* a *cardiaco* o *arteriosclerosis* a *arterioesclerosis*, mientras que otro cliente podría preferir el uso contrario (o incluso diferente) para los mismos términos. Todos ellos son correctos por diferentes motivos, por lo que los correctores ortográficos

* Este artículo está basado en un trabajo que se presentó en el I Congreso Internacional sobre Lenguaje y Asistencia Sanitaria, celebrado en Alicante en octubre de 2007.

** Universidad de Málaga (España). Dirección para correspondencia: claros@uma.es.

no los detectarán como erróneos o desconocidos y dependemos exclusivamente de nuestra habilidad.

- 2) Los clientes tienen usos muy particulares (incluso raros), como, por ejemplo, el de *abstracto* en lugar de *resumen*, *ribosomal* en lugar de *ribosómico* o «el enzima» en lugar de «la enzima». Aquí, los correctores ortográficos nos indicarían que los términos del cliente son erróneos, cuando sabemos que en ese caso no deben considerarse así. Esta situación induce a que no hagamos caso al corrector, con lo que el riesgo de dejar como correcto un término incorrecto es altísimo.
- 3) Sospechamos que en el texto van a aparecer términos que queremos evitar. Por ejemplo, queremos cambiar *clonaje* y *clonado* por *clonación*, *craneano* o *cranial* por *craneal*, *asumir* por *suponer*, *efectividad* por *eficacia*, *desfosforilar* por *desfosforilar*, *epinefrina* por *adrenalina*, etc. Los correctores ortográficos no nos van a ayudar, porque marcarán tanto los términos que hay que cambiar como los que ya hemos cambiado (*cranial*, *clonaje*, *clonación*, *desfosforilar*, *desfosforilar...*), o incluso darán por bueno algunos de los que queremos cambiar (*epinefrina*, *efectividad*, *asumir*).
- 4) Queremos estar seguros de las mayúsculas y minúsculas en algunos nombres que unas veces se escriben con mayúscula (por ejemplo, *mechero Bunsen*, *efecto Doppler* o *tubo Falcon*) y otras con minúscula (*un bunsen*, *el doppler* o *un falcon*), por lo que no hay manera de que el corrector ortográfico acierte en ambos casos.

Espero que nadie considere como solución práctica «la fuerza bruta», o sea, ponerse a buscar y reemplazar cada uno de los términos que aparecen en las directrices del cliente para comprobar si se han seguido correctamente. Además, somos humanos, por lo que, por más cuidado que pongamos, nunca estaremos seguros de que no se nos haya pasado algo y no queden términos que no son del gusto del cliente.

Por eso, creo que algo podremos aprovechar de esta humilde contribución al control terminológico de un texto. Como mínimo, debemos descubrir 1) que se puede tener un diccionario que corrija palabras aparentemente correctas en español, pero que son incorrectas en determinados casos o usos; 2) que este diccionario, además, puede llegar a sugerir la alternativa correcta al término señalado, y 3) que se pueden mantener varias listas negras diferentes y activar la(s) necesaria(s) para cada documento.

2. Las listas negras en Wordfast

Los programas para la traducción asistida (SDL Trados, Wordfast, Star Transit, Déjà Vu, etc.) suelen ofrecer un control de calidad para verificar que se han traducido correctamente los términos, de acuerdo con el glosario del proyecto o los que el traductor haya usado durante la traducción. Entre estos programas, el que realmente nos puede interesar es Wordfast, que ofrece además la posibilidad de crear una «lista negra» de términos (denominada «blacklist», figura 1A) que comprueba que no se dé por traducido un segmento

que contenga un término de dicha lista. Cuando Wordfast detecta un término de la lista negra en un segmento, indica al usuario la palabra o sintagma que debería verificar antes de darlo por traducido (figura 2B). Esta lista negra tiene la ventaja de que resulta fácil de crear —no es más que una lista de palabras separadas por retornos de carro— y tiene la gran ventaja de poder introducir tanto palabras únicas como sintagmas —en los que las palabras aparecen separadas por espacios o signos de puntuación, como «vida media» y «de lo anteriormente expresado se deduce que»—, e incluso reconocimientos difusos de las terminaciones, que se indican con comodines (*) como en «defosforil*», «glicosil*» o «hidrofóbic*». Pero también tiene algunos defectos: solo indica que hay una palabra que conviene revisar, sin proponer alternativas a ella. De la imagen de la figura 1A también se puede deducir que la lista negra se puede seleccionar o no, según sea necesario.

La principal limitación de la lista negra de Wordfast reside en que no se puede utilizar para una revisión de calidad de un texto que no sea bilingüe o no se haya escrito con MS Word —Wordfast solo funciona como un programa incrustado en él—. También es una pena que no sugiera alternativas cuando se encuentra un término «negro». Un último inconveniente de esta herramienta reside en que no gestiona bien las listas negras demasiado largas (más de 80 a 100 términos).

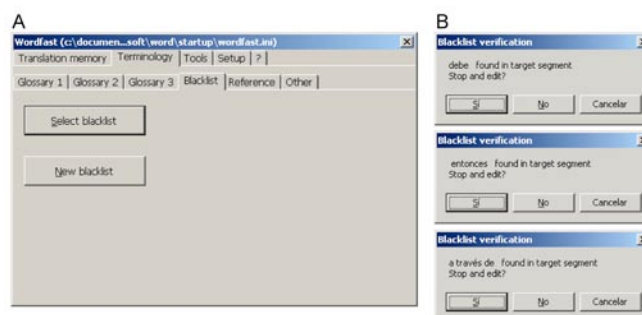


Figura 1. Distintos aspectos de la lista negra de Wordfast. **A:** Diálogo para incorporar una lista negra (*blacklist*) de palabras. Cuando se pulsa el botón «New blacklist», MS Word crea un documento alimentado con seis palabras en inglés para que el usuario lo use como modelo al incluir luego sus términos. **B:** Ejemplos de diálogos en los que se informa de que el párrafo traducido contiene un término o un sintagma erróneo o que conviene revisar

3. Listas negras en OpenOffice

Antes de comenzar este apartado, quiero que quede claro que, aunque OpenOffice podría sustituir a algunas de las herramientas que se usan con más frecuencia, no hay por qué cambiar nuestros hábitos, y simplemente basta con añadir OpenOffice a las herramientas de uso habitual siempre que el traductor, el escritor o el revisor lo desee. OpenOffice (figura 2) es un programa de dominio público y código abierto equivalente a Microsoft Office y totalmente compatible con él —abre y guarda documentos de Microsoft Office sin problemas, por lo que no hay que guardar documentos intermedios con cambio de formato—, hasta tal punto que da muchos menos problemas intercambiar documentos entre

Microsoft Office y OpenOffice, en cualquier sistema operativo, que intercambiar los documentos creados por Microsoft Office entre los sistemas operativos Windows y Mac OS X. A diferencia también de Microsoft Office, OpenOffice está disponible para los tres sistemas operativos más usados (Windows, Mac OS X y Linux), con total compatibilidad entre ellos; esto es, un documento creado con el OpenOffice de Windows se verá exactamente igual en el OpenOffice de los otros sistemas operativos, y viceversa (¿se atreve alguien a afirmar lo mismo respecto a los documentos de MS Office intercambiados entre Windows y Mac OS X?). Finalmente, el formato que utiliza para intercambiar sus archivos es el OpenDocument, que acaba de adoptarse como formato estándar de intercambio de documentos de texto a todos los niveles por la ISO/IEC 26300:2006 de 25/10/2006, que se puede consultar en <www.oasis-open.org/news/oasis_news_05_23_05.php>, por lo que es de esperar que pronto los organismos oficiales utilicen y acepten documentos en el formato OpenDocument. También conviene saber que ya se han escrito conversores para que Microsoft Office pueda leer y escribir en formato OpenDocument, aunque no vienen instalados por defecto. A título informativo, el formato PDF también ha sido recientemente aceptado como estándar según la norma ISO 32000-1, desde el 2/7/2008 (<www.iso.org/iso/pressrelease.htm?refid=Ref1141>). Resulta cuando menos curioso que el gigante Microsoft no incluya conversores para ninguno de los dos tipos de documento.

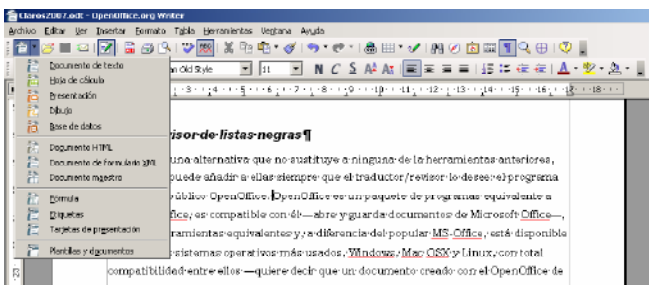


Figura 2. Aspecto de una ventana de OpenOffice en la que está desplegado el menú que muestra todos los tipos de documentos que se pueden abrir o crear

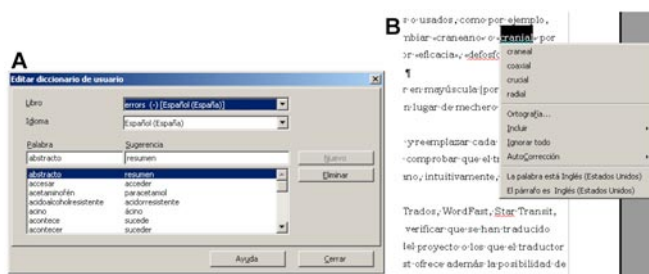


Figura 3. A: Ventana en la que se pueden introducir nuevos términos en la lista negra que contiene el diccionario «errors» para español, modificar los términos existentes o borrarlos. **B:** Menú contextual que aparece durante la edición del texto al pulsar el botón secundario en el que aparecen las alternativas ortográficas al término subrayado, en lo que aparece sobre el que estaba el cursor

Al intentar parecerse a Microsoft Office, OpenOffice contiene herramientas equivalentes a las de aquel, como la de corrección ortográfica en numerosos idiomas y los diccionarios personales, gracias a lo cual se subrayan los términos erróneos del texto. Pero también presenta herramientas nuevas, como la capacidad de definir una lista negra de palabras y su corrección (figura 3A). Gracias a ella, también se subrayan en pantalla las palabras que aparecen en la lista negra. Cuando se pulsa el botón secundario del ratón sobre la palabra subrayada (figura 3B), aparecen las alternativas «correctas».

4. Creación de la lista negra en OpenOffice

Para crear una lista negra en OpenOffice hay que acceder al menú «Herramientas» > «Opciones...» (figura 4A) para que aparezca la ventana de diálogo de la figura 4B. Entonces se escoge, en el marco que contiene el menú expandible de la izquierda, el detalle «Configuración de idioma» > «Lingüística». Veremos que la parte de la derecha ha cambiado con la selección, de manera que ahora se puede pulsar el botón «Nuevo...» para pasar a la configuración del nuevo diccionario en otra ventana de diálogo, titulada «Crear diccionario de usuario» (figura 4C). En dicha ventana hay que dar un «Nombre» al diccionario, escoger en el menú desplegable el «Idioma» con el que se quiere emplear y, muy importante, marcar la casilla correspondiente a «Excepción (-)», que es la que hará que dicho diccionario sea una lista negra en lugar de un diccionario personal como todos los demás. Una vez realizadas las tres acciones (dar nombre, escoger idioma y marcar que será una excepción), se pulsa el botón «Aceptar» (figura 4C), y el nuevo diccionario queda definido y aparece en la franja central de la ventana de diálogo titulada «Opciones – Configuración de idioma – Lingüística» con la calificación de «(-)» y la indicación del idioma para el que funcionará (figura 4B).

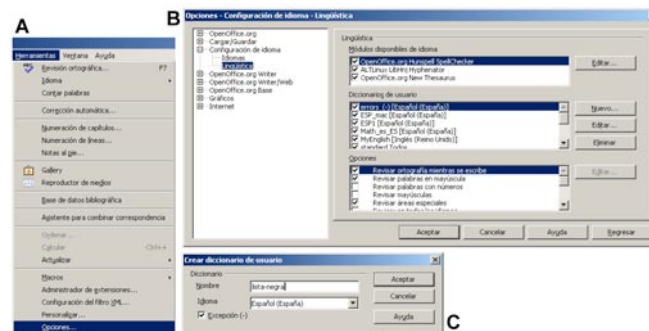


Figura 4. Imagen de las tres etapas necesarias (A, B y C) para crear un diccionario que contendrá la lista negra de palabras. Aunque en la ventana C se ha denominado *lista-negra* al diccionario, a lo largo de este texto se usará el diccionario «errors», que aparece el primero en la lista de la ventana de la imagen B

Para comenzar a introducir términos, hay que seleccionar el diccionario de excepciones (llamado «errors» en la figura 4B) y pulsar el botón «Editar...». Así se abre una nueva ventana de diálogo (como la que aparece en la figura 3A), en la que se puede escribir el término considerado incorrecto en

campo «Palabra» y el término correcto en el campo «Sugerencia». Podemos desde no usar ningún diccionario de listas negras hasta emplear varios a la vez —por ejemplo, uno por cliente—, ya que se pueden activar y desactivar a voluntad en cualquier momento. Para ello basta con hacer clic en la casilla que aparece a la izquierda del nombre del diccionario (figura 4B). Por desgracia, este diccionario de excepciones tiene la limitación de que no se pueden incluir reconocimientos difusos, ni sintagmas, ni tampoco signos de puntuación o espacios, ya que ambos los utiliza OpenOffice para delimitar las palabras.

5. Conclusión

Si se usan las listas negras, es factible un esquema de trabajo cómodo y eficaz que permita que la misma persona traduzca y revise un texto para cotejar la coherencia terminológica de forma fiable. El procedimiento sería el siguiente:

1. El texto en español se escribe a gusto del redactor o revisor y se guarda en un fichero de MS Word, o bien en formato OpenDocument, RTF, HTML o texto simple.
2. El revisor, tanto si es la misma persona que escribió el texto como si es otra distinta, abre el fichero generado en 1) con OpenOffice. Es importante insistir en que no se necesita ningún tipo de conversión de formatos.
3. El documento abierto en OpenOffice mostrará todas las palabras desconocidas subrayadas en rojo. Entonces se pulsa el botón secundario del ratón para que aparezca el menú contextual (figura 3B), en el que se indican, entre otras cosas, las correcciones sugeridas. No se necesita abrir ninguna ventana adicional ni pulsar botones de «Ignorar», ya que el revisor es quien decide qué palabras podrían necesitar una corrección. Por si no estaba claro, las palabras subrayadas son tanto las que presentan errores ortotipográficos como las de las listas negras.
4. Una vez revisado todo el documento, este se guarda en el mismo formato en que estaba, simplemente escogiendo el menú «Archivo» > «Guardar».

Por tanto, para revisar un texto con OpenOffice, que pueden usar y compartir el traductor y el revisor si lo desean, solo necesitan tener dicho programa instalado en su ordenador —se puede descargar de forma gratuita desde <www.openoffice.org>; la instalación es muy sencilla—, haber definido su propio diccionario de errores y usar prácticamente cualquier fichero que proceda de un procesador de textos o del paquete Microsoft Office. Para aquellos que quieran ponerlo en práctica, el autor pone a su disposición su diccionario de términos que no le gustan en español (el diccionario «errors» que aparece en la figura 4B), que ha de colocarse en la carpeta de Windows C:\Documents and Settings\«nombre del usuario»\Datos de programa\OpenOffice.org2\user\wordbook\ (se trata de una carpeta invisible, porque la carpeta «Datos de

programa» lo es) o, en el Mac OS X, dentro de la carpeta Librería de la cuenta del usuario ~Librería\Application Support\OpenOffice.org2\user\wordbook\. De aquí se deduce que el mismo fichero de diccionario vale para todas las versiones de OpenOffice que existen, sin necesidad de conversión.

Aquellos que quieran hacer visible la carpeta «Datos de programa» de Windows no tienen más que ir a la carpeta C:\Documents and Settings\ y escoger el menú «Herramientas» > «Opciones de carpeta» (figura 5A). Aparecerá una ventana en la que hay que escoger la pestaña «Ver» (figura 5B). Nos deslizaremos hasta encontrar la línea que dice «Archivos y carpetas ocultos», donde seleccionaremos el botón radial «Mostrar todos los archivos y carpetas ocultos» (figura 5B). Luego cerraremos la ventana y observaremos que todas las carpetas que antes no veíamos ahora se ven atenuadas (figura 5C) y se puede navegar por ellas.

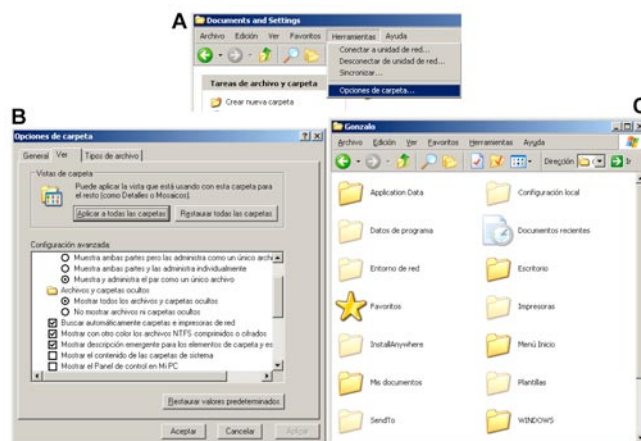


Figura 5. Etapas que hay que seguir para hacer visibles las carpetas ocultas, de manera que podamos añadir o quitar ficheros manualmente en el directorio donde están los diccionarios de OpenOffice. **A:** La selección de la opción de menú «Opciones de carpeta...» abre el diálogo **B**, donde hay que elegir la opción «Archivos y carpetas ocultos» > «Mostrar todos los archivos y carpetas ocultos», que hace visibles de forma atenuada (**C**) las carpetas y ficheros antes ocultos

A modo de primicia, puedo anunciar que estoy a punto de permitir el uso público de un programa en Java —de manera que se pueda utilizar en todos los sistemas operativos tal cual— para manejar listas negras de palabras y sintagmas, junto con el motivo por el que se considera incorrecto el término seleccionado. La idea es que el redactor, el traductor y el revisor puedan crear y modificar estas listas con total libertad y de forma sencilla, añadiéndoles términos, sintagmas y caracteres comodín, además de una explicación que ayude a decidir si el cambio se debe realizar o no. De esta forma, los revisores y redactores elegirán con facilidad las listas negras que quieren utilizar con cada cliente, y las agencias y editoriales contarán con una herramienta que les facilitará la revisión de estilo y la coherencia terminológica. Pero, mientras tanto, las listas negras con OpenOffice son un buen sustituto.