FRAGMENTOS LÉXICOS DE LA TRADUCCIÓN DE TEXTOS

María Antonia Babí i Vila

University of Swansea (Reino Unido)

Resumen

Esta comunicación propone un nuevo enfoque para analizar traducciones y detectar plagios. Mediante la comparación de fragmentos léxicos de un determinado número de palabras se puede averiguar si éstos son idénticos. Los resultados apuntan a la superposición de fragmentos cortos similares.

El análisis se realizó entre las 100 primeras palabras de cincuenta traducciones del español al inglés de un texto periodístico. El escrutinio detallado basado en fragmentos léxicos de cinco, seis y siete palabras es significativo. Aunque hay gran parecido entre dichos fragmentos es difícil detectar la igualdad. Los fragmentos de más de 20 palabras son raros. La presente comunicación sugiere que los casos de supuesto plagio pueden evaluarse siguiendo modelos informáticos.

1. INTRODUCCIÓN

En esta comunicación me propongo explorar un problema de evaluación real que se produjo en un examen internacional de nivel avanzado. El problema surgió cuando dos de los candidatos a un reconocido diploma de traducción de español al inglés, presentaron dos traducciones casi idénticas.

Creo que a todos los profesores nos ha pasado cuando nos hemos encontrado delante de un fajo de exámenes para corregir. Al empezar la tarea cogemos uno y le asignamos una nota, pero hacia la mitad o al final del fajo nos aparece otro que pensamos que ha sido copiado de uno que ya habíamos corregido antes. En primer lugar, la reacción emotiva de incredulidad

Centro Virtual Cervantes

nos asalta ¿Se trata de un espejismo? Y en segundo lugar ¿qué hacemos ante el fenómeno? Hoy en día las políticas educativas no nos permiten eludir el problema. En realidad, y me estoy refiriendo a niveles avanzados, el plagio[1] no sólo se penaliza duramente sino que se está convirtiendo en una actividad muy compleja a la que personas de gran experiencia docente dedican parte de su actividad profesional. En este caso, me refiero a la situación universitaria de Gran Bretaña y del mundo anglosajón. En concreto, a finales de curso, mi institución nos dio unos cursos de preparación tecnología explicándonos como detectar el plagio mediante programas de ordenador. Así pues, y a pesar de los pesares, debemos lanzarnos a descubrir nuevas tecnologías que faciliten nuestro trabajo y que nos ayuden a presentar los procesos de corrección con rigor científico y claridad. Debo aclarar que los exámenes de lengua son aún más difíciles de detectar que los exámenes de contenido de literatura o cultura.

El texto original consistía en una reseña, específicamente, trataba de la descripción de uno de los festivales de cine de Canes que apareció en la revista semanal de *El País*. Para conseguir el diploma de traducción los candidatos debían presentar un texto completo con anotaciones extras comentando las peculiaridades del texto original y sus dificultades. El análisis que presento intenta descubrir cómo es posible justificar que dos de los textos no fueron copiados.

Los candidatos eran de procedencia internacional y principalmente europeos que se presentaban a un examen de postgrado. Las exigencias del ejercicio eran muy estrictas pero se permitía a los candidatos usar un diccionario. Tal vez este hecho explicaría que en los textos hubiera palabras poco frecuentes para hablantes de segunda lengua.

2. DESCRIPCIÓN DEL MATERIAL

El análisis va más allá de la anécdota e intenta justificar el fenómeno con un método empírico de tecnología avanzada. Originariamente trabajamos con 119 traducciones pero hubo que descartar algunos por no estar completos. Cada texto se transcribió en formato ASCII y el programa operativo que se manejó para analizarlos fue el *V-tools*. Al final, concentramos el análisis en cincuenta traducciones ya que los resultados parecían ya representativos y se calcularon dos conceptos: el número total de palabras escritas y, el número

[1] Para la definición de plagio ver Alexander Lindsay (1952: 2) *'Plaigarism and Originality',* publicado por Harper, Nueva York.

Centro Virtual Cervantes

de palabras que se repetían para cada texto. Las traducciones al inglés, como se ilustra a continuación en el cuadro 1 son similares, sin embargo no son idénticas. La variedad entre ellas se manifiesta sobre todo, o bien en la longitud del texto, o bien en el número de palabras. Concretamente, cada texto consta de un total de 500 a 750 palabras *(tokens)* y entre ellas hay de 275 a 375 palabras únicas *(types)*.



Gráfico 1. Resultado global de las palabras traducidas.

3. CONTENIDO DE LAS TRADUCCIONES

Para el presente análisis me limito al primer párrafo de la traducción. Dicho párrafo en la versión original en español, consta de una oración de unas cincuenta palabras.

A continuación en el cuadro 1, se presentan algunas versiones de candidatos en inglés.

Nótese que si bien son similares, no son ni mucho menos idénticas.

Pero no olvidemos que nuestro objetivo es la comparación entre textos y la posibilidad de que dos de los candidatos hayan intercambiado opiniones, ya sea por medio de tecnología, o por gestos, o simplemente ayudándose mutuamente. En el siguiente cuadro y gráfico 2, aparece la lista de todas las palabras que se repiten en todos los textos. Exactamente, se trata de 30 vocablos. A excepción de los sustantivos propios, son las palabras típicas, artículos, pronombres, verbos, como no, el frecuente ser, preposiciones, conjunciones y algún sustantivo común como cine o festival.

Centro Virtual Cervantes

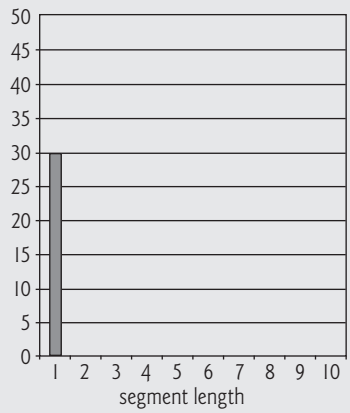
ANÁLISIS DE LOS TEXTOS

- The enormous, precise, perfect and because of its impenetrability, cruel machinery of the Cannes Film Festival...
- The Cannes Film Festival machine, huge, precise, perfect and harsch in its all-encompassing nature...
- The enormous, precise, perfect and, judging by its most clearly distinguishable characteristics, cruel machine of Cannes Festival...
- It is almost upon us, the enormous, precise, perfect and for those that it excludes, cruel machinery of the Cannes Festival...
- Now the enormous, precise, perfect and judging by its vastness, cruel machinery of the Cannes Film Festival is about to stir...
- It is almost in motion –the enormous, precise, perfect and, because of its extensive nature, cruel machinery of the Cannes Film Festival...
- The enormous, precise, perfect extensive and cruel machinery of the Cannes Film Festival is about to be set in motion...
- It is already beginning to stir; the huge, defined flawless, but most of all cruel machinery that is the Cannes Film Festival...
- So the vast, precise, well oiled and, due to its unassailability, cruel machinery of the Cannes Festival is almost on the move...

Cuadro 1. Ejemplos de traducciones del primer párrafo del artículo.

PALABRAS QUE APARECEN EN TODOS LOS TEXTOS

a	an
and	are
be	but
by	Cannes
festival	festivals
film	for
in	is
it	Julia-Roberts
like	not
of	or
that	the
their	they
this	to
who	will
with	Woody-Allen

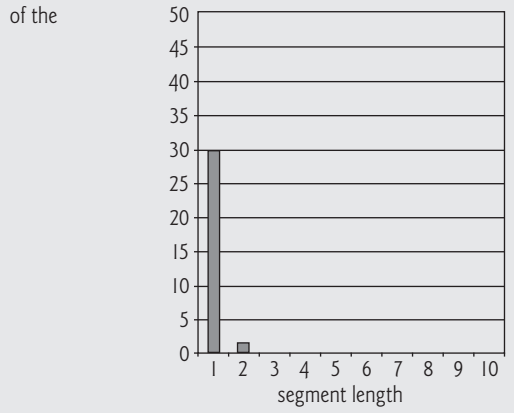


Cuadro y gráfico 2. Palabras comunes a los 50 textos.

Centro Virtual Cervantes

La sorpresa aparece al analizar qué ocurre con fragmentos de texto de dos palabras, en el siguiente cuadro y gráfico, el número 3.

SERIE DE DOS PALABRAS QUE APARECEN EN TODOS LOS TEXTOS



Cuadro y gráfico 3. Serie de 2 palabras.

La búsqueda se hace más interesante si comparamos dos textos escogidos al azar, y además analizamos fragmentos de seis o más palabras. Los patrones se producen pero muy aisladamente. Es decir, estamos muy lejos de conseguir unas traducciones homólogas, la variedad parece la norma.

FRAGMENTOS DE 6 PALABRAS QUE APARECEN EN DOS TEXTOS ESCOGIDOS

...12 days of cinema and roses...
...by an army of famous faces...
...the four corners of the planet...
...to a severe type of cinema...
...godfather impossible to reject savage and...
...it without giving anything in return...
...language of film so that it...
...and despotic the fact that Cannes...

...creation of art by those who like...
...the enormous precise perfect and because of its...

Cuadro 4. Fragmentos léxicos de 6 o más palabras.

Centro Virtual Cervantes

FRAGMENTOS DE N PALABRAS QUE APARECEN EN DOS TEXTOS ESCOGIDOS

7 palabras
creation of art by those who like...

8 palabras
the enormous precise perfect and because of its...

10 palabras
To create mutations in the eloquence of the language of...
the competition within the enormity of Cannes is small but...

11 palabras
About the fruits of the harvest of what the organisers of...
The admiration and unconditional support of any star who under the...

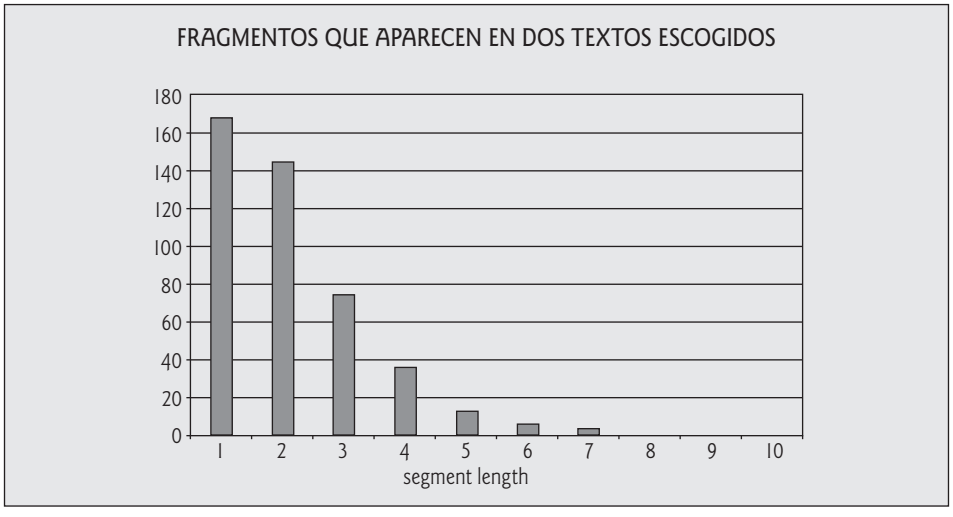
Cuadro 5. Fragmentos léxicos de más de 6 palabras.

FRAGMENTOS DE N PALABRAS QUE APARECEN EN DOS TEXTOS ESCOGIDOS

23 palabras
...to the despotism of a manipulated market is equivalent to winning in one fell swoop the admiration and unconditional support of any star...

Cuadro 6. Fragmentos léxicos de más de 20 palabras.

La siguiente gráfica, la 4, ilustra el total de fragmentos léxicos comunes que aparecen en dos de los textos agrupados según el número de palabras.



Gráfica 4.

Centro Virtual Cervantes

4. CONCLUSIÓN

La primera observación es que, aunque no es imposible encontrarlas, hay pocas series de veinte palabras. Es más, se encuentran muy pocas series de más de veinte palabras. Es decir, los candidatos actuaron contrariamente al texto original. Por lo tanto la posibilidad de que existan dos textos compartiendo segmentos de cincuenta palabras es muy remota.

Segunda observación, curiosamente y como ya se habrá deducido por los ejemplos que se han ilustrado anteriormente, los segmentos de diferentes números de palabras no corresponden a unidades sintácticas del texto original. La peculiaridad aquí es la falta de estructuras características a la clasificación del español a oraciones ya sean simples o complejas. Cabría la posibilidad de que los resultados expuestos podrían deberse al estilo del texto original. Éste contiene oraciones muy largas, con secuencias de varios adjetivos y clichés.

Tercera observación, El tribunal de examinadores, y después de llevadas a cabo las investigaciones oportunas, concluyó que las dos traducciones no habían sido copiadas. Las razones por las cuales se llegó a dicha decisión fueron varias. Una de ellas fue la distinta procedencia de los candidatos y la edad: se trataba de un joven inglés y una alemana jubilada y además, se dio la casualidad que habían estado sentados en extremos opuestos del aula de examen.

Finalmente, cabría pensar en las implicaciones legales que podrían afectar tanto a los estudiantes que pudieran ser acusados de plagio como a las instituciones cuyos profesores han tomado las decisiones que lo propusieran, pero mi intención en la comunicación ha sido dar nueva luz a un viejo problema. El trabajo está por hacer.

AGRADECIMIENTOS

- A los amigos de ASELE y de Cáceres.
- Prof. P. M. Meara que ha facilitado el análisis de los textos (programa V_Tools).
- <http://www.lognostics.co.uk/>.
- Al Institute of Linguists de Londres que nos ha dado permiso para usar sus exámenes.
- <m.a.babi@swansea.ac.uk>.
- Swansea University.

Centro Virtual Cervantes