

# **Cómo investigar en Traducción Audiovisual: archivos, metodología y herramientas<sup>1</sup>**

**Isabel Briales Bellón**

Universidad Pablo de Olavide. Facultad de Humanidades. Departamento de Filología y Traducción. Sevilla, España

[ibribel@upo.es](mailto:ibribel@upo.es)

## **Research in Audiovisual Translation: files, methodology and tools**

Fecha de recepción: 22.10.2021 / Fecha de aceptación: 17.12.2021

*Tonos Digital*, 42, 2022

### **RESUMEN:**

Corpas Pastor (2008: 89) señala que la Lingüística de Corpus está considerada como un enfoque metodológico adecuado para el estudio de las lenguas, aplicable tanto a la traducción como a la interpretación, pues permite analizar y describir una amplia variedad de discursos, reflexionar sobre la unidad de traducción y revisar el concepto de equivalencia. Partiendo de la base de que es factible trabajar con *corpus* en ámbitos tan dispares como la investigación sobre la variación lingüística o durante la realización de un encargo profesional o el desarrollo de una metodología para la enseñanza de lenguas, presentamos las fases que hemos tenido que superar para llevar a cabo el estudio comparativo del guion original del filme *Bienvenue chez les Ch'tis* y las versiones doblada y subtitulada en español. Destacamos la importancia de preparar correctamente el material de trabajo para poder extraer el máximo provecho de él. En el caso que presentamos, hemos utilizado los programas Final Draft, AntConc y WinAlign, que nos han permitido profundizar y desgranar los textos, localizar fácilmente los segmentos interesantes y crear unidades de traducción. Se trata de una metodología sistemática que permite abordar con rigurosidad la investigación en Traducción Audiovisual.

---

<sup>1</sup> La investigación realizada para el presente estudio forma parte de la tesis doctoral de la autora.

**Palabras clave:** Traducción Audiovisual; *corpus*; herramientas; metodología; unidad de traducción

**ABSTRACT:**

Corpas Pastor (2008: 89) notes that corpus linguistics is considered a valid methodological approach to language studies, relevant to both translation and interpreting insofar that it can be used to analyse and describe a wide range of discourse types, reflect on translation units and re-examine the concept of equivalence. Taking as its starting point the fact that it is possible to work with corpora in such varied undertakings as researching language variations, performing professional translation/interpreting tasks and developing language teaching methods, this paper describes the stages involved in comparing and contrasting the original script for the film *Bienvenue chez les Ch'tis* and its dubbed and subtitled renderings in Spanish. We highlight the importance of properly preparing subject material to be able to make the most of its potential. In this particular case, we used the Final Draft, AntConc and WinAlign programmes to analyse the texts in depth, easily identify segments of interest and create translation units: a systematic method which enables research into Audiovisual Translation to be carried out in a rigorous manner.

**Keywords:** Audiovisual Translation; corpora; tools; methodology; translation unit

## **1. INTRODUCCIÓN Y OBJETIVO DEL ESTUDIO**

*Bienvenue chez les Ch'tis* es un filme francés que apareció en las pantallas del país galo en 2008 y en España en 2009, con el fin de mostrar al mundo los prejuicios y las diferencias entre el norte y el sur de Francia. La razón por la que esta película captó nuestra atención desde el principio es la siguiente: el hilo conductor está basado en la variedad lingüística, con la presencia constante del dialecto *ch'timi*, en el humor y en las referencias culturales, lo que dificulta en gran medida la labor del traductor.

Este artículo tiene como objetivo mostrar las fases de preparación y tratamiento del archivo textual formado a partir del guion original del filme francés *Bienvenue chez les Ch'tis* y las versiones doblada y subtitulada en español. El objetivo final de nuestro trabajo, para el que preparamos el archivo textual, sería analizar en profundidad estos guiones con la intención de identificar las unidades léxicas marcadas por la variedad lingüística y, a partir de ahí, las estrategias empleadas en el trasvase de este contenido presente en *Bienvenue chez les Ch'tis* a la lengua española (*Bienvenidos al Norte*). Como hemos mencionado, pretendemos investigar

este tema a partir de un archivo textual definido, enlazando aspectos lingüísticos y culturales que no pueden ser considerados compartimentos estancos en la labor de traducción.

Para ello, utilizaremos tres herramientas concretas: FinalDraft, AntConc y WinAlign. Hemos seleccionado estos tres programas porque, desde nuestro punto de vista, son idóneos para acceder al contenido de los archivos originales, desmenuzar el texto y localizar en él las unidades léxicas provenientes del *ch'imi*, así como ubicar los equivalentes en español presentes en las versiones doblada y subtitulada. Así, FinalDraft nos permitirá adentrarnos en el guion de preproducción elaborado por Dany Boon, que ha sido creado con esta herramienta y tiene una extensión *fd*; gracias a AntConc localizaremos las unidades léxicas que nos interesan, ya que nos mostrará en su contexto todas y cada una de las veces que cada unidad seleccionada aparece en el guion; y, por último, WinAlign es la herramienta ideal para alinear original y traducción, es decir, para emparejar los segmentos localizados en el paso anterior con sus equivalentes en los guiones de doblaje y subtulado.

## **2. CORPUS VERSUS ARCHIVO TEXTUAL**

Corpas Pastor (2008) hace hincapié en la importancia y versatilidad de la Lingüística de Corpus, ya que está considerada como un enfoque metodológico adecuado para el estudio de las lenguas, que se puede aplicar tanto a la traducción como a la interpretación, habida cuenta de que permite analizar y describir una amplia variedad de discursos, reflexionar sobre la unidad de traducción y revisar el concepto de equivalencia.

De igual forma, la investigación basada en *corpus*<sup>2</sup> constituye los cimientos de los programas de traducción automática, de extracción terminológica o es el componente esencial de los estudios contrastivos, siendo los dos tipos de *corpus* más empleados para estos fines los comparables y los paralelos.

*En la lingüística de corpus se diferencia entre textos comparables en la lengua meta (esto es, textos originales en la lengua meta producidos en situaciones comunicativas similares, que versan sobre el mismo tema que el texto origen y pertenecen al mismo tipo textual) y textos paralelos, que comprenden el texto origen y su(s) traducción(es) a la lengua meta. Corpas Pastor (2008: 89)*

Asimismo, se utilizan todos los recursos que brinda la Lingüística de Corpus para indagar sobre variedades de una lengua, registros y uso, por lo que Corpas

---

<sup>2</sup>Emplearemos la forma invariable *corpus* tanto para el singular como para el plural, tal y como se indica en el Diccionario Panhispánico de Dudas (RAE, 2005).

Pastor aboga por aprovechar las posibilidades que ofrece un *corpus* tanto para profesionales del mundo de la traducción como para discentes y docentes de esta materia que deben enfrentarse a un encargo, real o no, de traducción.

Cabe destacar que la implantación de esta forma de trabajo no arranca desde cero, ya que tradicionalmente se han utilizado textos comparables como fuente de documentación, habida cuenta de que en numerosas ocasiones, son mucho más fiables que el propio diccionario.

Así, partiendo de la base de que es factible trabajar con *corpus* en ámbitos tan dispares como la investigación sobre la variación lingüística o durante la realización de un encargo profesional o el desarrollo de una metodología para la enseñanza de lenguas, nos disponemos a presentar el material seleccionado, no sin antes profundizar en el concepto de *corpus*, en tanto que colección más o menos extensa de texto en soporte electrónico que debe ser representativo del estado de lengua o de la sección de lengua elegida. De la misma manera, recoge los criterios mínimos que establece Sinclair (*apud* Corpas Pastor, 2001: 157) para que las colecciones de texto en soporte informático puedan ser consideradas *corpus*, esto es, «cantidad (extensión del *corpus* en número de palabras), calidad (representatividad y equilibrio), simplicidad de codificación, y documentación (referencia exacta a las fuentes documentales utilizadas)».

De igual forma, Parodi (2010: 21) se hace eco de las recomendaciones del grupo EAGLES (*Expert Advisory Group on Language Engineering Standards*) en el marco de un proyecto de la Unión Europea para la estandarización de los trabajos que se llevan a cabo en las diversas lenguas europeas. Así pues, EAGLES propone las siguientes pautas para que un *corpus* pueda ser considerado como tal:

1. El corpus debe ser lo más extenso posible de acuerdo con las tecnologías disponibles de cada época.
2. Debe incluir ejemplos de amplia gama de materiales en función de ser lo más representativo posible.
3. Debe existir una clasificación intermedia en los géneros entre el corpus en total y las muestras individuales.
4. Las muestras deben ser de tamaños similares.
5. El corpus, como un todo, debe tener una procedencia clara.

Vistas las características que debe tener un *corpus*, podemos afirmar que los guiones que constituyen nuestro material de trabajo no conforman un *corpus* sino un

archivo textual, ya que según Corpas Pastor (2008: 145), la definición de este último y las diferencias con el *corpus* son las siguientes:

**archive** Generally defined as being similar to a **corpus**, although with some significant differences. Geoffrey Leech (1991: 11) suggests that 'the difference between an archive and a corpus must be that the latter is designed for a particular "representative" function'. An archive, on the other hand, is simply 'a text repository, often huge and opportunistically collected, and normally not structured' (Kennedy, 1998: 4) (Baker, Hardie y McEnery, 2006).

La elaboración del *corpus* está planificada de forma clara y concisa, mientras que los archivos textuales son colecciones de texto en formato electrónico, que pueden abarcar desde bibliotecas digitalizadas hasta documentos obtenidos de Internet. El archivo textual, objeto de estudio de este trabajo, ha sido compilado siguiendo la metodología de compilación de *corpus ad hoc* virtual descrito en Corpas Pastor (2008), en donde se denomina *corpus ad hoc* a un *corpus* virtual compilado con carácter puntual para la realización de un encargo de traducción determinado, independientemente de la dirección (directa, inversa o indirecta), siendo el aspecto más importante la calidad de los textos que lo conforman.

Normalmente los *corpus ad hoc* no están constituidos por un gran número de textos, sino por textos muy representativos, equiparables al original en cuanto a temática, género y variedad textual. El principal objetivo de este tipo de *corpus* es la creación de un recurso documental de fácil obtención, fiable y específico, dedicado a resolver problemas de reexpresión, interpretación del discurso y toma de decisiones. La compilación de *corpus ad hoc* requiere seguir un protocolo de análisis del texto origen con vistas a su traducción, selección de los sistemas de recuperación de información, evaluación de las fuentes y de los recursos electrónicos y comprobación del grado de correspondencia entre el *corpus* y el texto original, atendiendo a características lingüísticas y textuales.

Conviene insistir en que el *corpus ad hoc* es un caso claro de *corpus* virtual y, de hecho, esta característica forma parte de su definición. La principal diferencia entre *corpus* estable y virtual es que este último presenta un diseño puntual, condicionado por una finalidad específica y por utilizar métodos de compilación alternativos que no incluyen fase de escaneo ni de anotación, puesto que habitualmente los textos se encuentran en formato digital. Por el contrario, un *corpus* estable está sujeto a un diseño y tamaño fijo, aunque adaptable, y cuya elaboración pasa por el escaneo de textos, con quizás anotación interna y métodos de recuperación de información

propios. Un *corpus* que originalmente surgió como virtual puede llegar a ser estable si alcanza un nivel de institucionalización mínimo.

Una vez revisados los conceptos teóricos, nos centramos en el archivo textual seleccionado para realizar este estudio, que está compuesto por tres guiones: uno en francés, escrito por Dany Boon y al que podemos denominar 'guion original', y dos provenientes de la traducción al español de la película en francés, uno procedente del doblaje y otro del subtulado de dicho filme.

El archivo textual, teniendo en cuenta que nuestro objetivo es el análisis de las versiones francesa y española, es representativo porque está constituido por la transcripción completa de los diálogos de pantalla en ambas lenguas, así como la transcripción del subtulado en español. Hemos excluido la transcripción de los subtítulos en lengua francesa porque no es relevante en este estudio.

En el caso que nos ocupa, se trata de un archivo textual extenso en cuanto a número de palabras, pero reducido, si nos fijamos en el número de textos incluidos, pues no existe ni otra versión original de *Bienvenue chez les Ch'utis* ni otra traducción al español realizada en otro momento o con otra variedad que podamos obtener para nutrir nuestro material de trabajo, por lo que se limita por sí solo en cuanto a aspectos temporales y geográficos se refiere. Cabe señalar que el filme fue exportado a Argentina bajo el título *Bienvenidos al país de la locura*, pero la variedad empleada para el doblaje es la del español de España, es decir, que el título es lo único que diferencia a esta película de *Bienvenidos al Norte*.

Podríamos decir que se trata de un archivo textual conformado por textos paralelos, según el tipo de texto que contiene y siguiendo la clasificación para *corpus* pero aplicable a este tipo de compilación, ya que se ajusta a la definición que se propone:

*Corpus paralelo. Corpus formado por una serie de textos en la lengua origen junto con sus traducciones en una (o varias) lengua(s) meta. Se habla de corpus paralelo bilingüe cuando hay solo dos lenguas implicadas, y de corpus paralelo multilingüe cuando se trata de más de dos. Corpas Pastor (2001: 158)*

La principal diferencia entre '*corpus* paralelo' y '*corpus* comparable' radica en que el *corpus* comparable está formado solo por textos originales en una determinada lengua, escritos en situaciones comunicativas parecidas a las de otros *corpus* de otras lenguas y que, siendo el mismo género textual, responden a idéntica temática. Al ser textos equiparables por su similitud en forma y función, es posible la comparación interlingüística de los elementos.

### **3. GUIONES DE PREPRODUCCIÓN Y DE POSTPRODUCCIÓN: EL ARCHIVO TEXTUAL**

Para llevar a cabo nuestra labor de investigación ha sido muy importante contar con el guion original en francés, elemento clave del archivo textual. La película contiene léxico y expresiones en *ch'timi* cuya transcripción no es fácil, por lo que, antes de realizar tan ardua tarea nosotros mismos, nos pusimos en contacto con la productora de *Bienvenue chez les Ch'tis*, Pathé, que muy amablemente y en un brevísimo espacio de tiempo nos facilitó el guion original completo. Nos hicieron llegar un archivo denominado *BienvenuechezlesCh'tisV3du13-4-2007.fdr*, que indicaba que no era el guion definitivo sino que se trataba de la tercera versión (V3), que fue terminada el 13 de abril de 2007 y que tiene una extensión .fdr. En un primer momento, desconocíamos qué programa debíamos adquirir para poder abrir archivos .fdr. Tras indagar, descubrimos que se trataba de un documento elaborado con Final Draft, *software* específicamente diseñado para la creación de guiones.

Al ser el guion original, o siguiendo la terminología propuesta por Chaume (2004), 'guion de preproducción', hubo que realizar una verificación de la lista de diálogos, es decir, comprobar si el guion se correspondía con el de postproducción (Chaume, 2004) o 'guion de pantalla', definido por Fuentes Luque (2000) como «lo que en realidad se dice en la pantalla, y que puede o no coincidir con el guion original».

El proceso, aunque lento, nos permitió ir anotando todas las diferencias existentes entre los guiones de preproducción y postproducción. El resultado es muy interesante, ya que en el guion de preproducción aparecen escenas y diálogos posteriormente eliminados o transformados y que nos permiten recrear la idea original de Dany Boon.

En relación con los guiones en español fruto de la traducción de *Bienvenuechez les Ch'tis*, nos pusimos en contacto con Wanda Visión, distribuidora de la película en España, para solicitar información sobre qué empresa había realizado la traducción al español del filme francés y de qué manera podíamos obtener los guiones de las versiones doblada y subtitulada. Al no recibir respuesta, tomamos la decisión de transcribir a mano los diálogos de pantalla, siendo conscientes de que en determinados momentos esta tediosa tarea podía tornarse complicada, puesto que en la versión doblada al español, el acento de los habitantes de Nord-Pas de Calais y la aparición de palabras inventadas en los diálogos dificultan a veces la comprensión.

Una vez recopilados los guiones que componen el archivo textual, comenzamos con las diferentes etapas del análisis.

En primer lugar, efectuamos un preanálisis de los archivos, identificando los aspectos a los que Berman (1995) alude en la descripción general de las pautas para estudiar traducciones, haciendo, sobre todo, hincapié en marcar las secuencias significativas en los filmes en las que abunda el léxico en *ch'timi*.

### **3.1. Berman: *Pour une critique des traductions***

Con el fin de abordar el estudio empírico, adoptamos como referencia la metodología descrita por Berman en *Pour une critique des traductions: John Donne* (1995), donde marca las pautas para el análisis de traducciones, que se puede adaptar al formato textual en el que luego aparecerá publicado el estudio, es decir, se puede adecuar la metodología al formato y a la longitud de la obra en la que vea la luz, pues el análisis de un caso no será igual de exhaustivo ni extenso en una comunicación de congreso, una tesis doctoral o un artículo de revista.

Del discurso recogido en dicha obra se desprende, además y gracias a los ejemplos que aporta, que Berman lleva a cabo la comparación entre obras literarias, por lo que adaptaremos las pautas al archivo textual con el que trabajamos, es decir, al guion de la película *Bienvenue chez les Ch'tis* y su traducción al español.

En cuanto al procedimiento, diríamos que, en primer lugar y siempre según Berman (1995), es necesario leer la traducción como si de un original se tratara, dejando a un lado por completo el texto de partida. De hecho, una vez se ha decidido llevar a cabo el estudio, recomienda abordar la lectura del texto meta (en adelante TM) antes que la del texto original, (en adelante TO), por la sencilla razón de que seamos capaces de mantener suficiente distancia con respecto al TO, intentando tener una visión del TM como material independiente.

Una vez hecho esto, aconseja volver a leer la traducción con un doble objetivo:

Por un lado, comprobar si se ajusta a las tipologías textuales existentes en la lengua meta, y, por otro, determinar si la redacción guarda cohesión y coherencia como texto individual, independientemente de la relación que mantenga con el documento original. En este momento, será posible marcar las expresiones, los términos o los fragmentos "problemáticos", que hacen que la traducción pierda fluidez.



Insiste Berman en la importancia de dejarse llevar por las "impresiones", por las sensaciones que la traducción nos transmite, pues, finalizado el primer acercamiento al texto meta, nos enfrentaremos al original, y dichas impresiones, que después se revelarán erróneas o acertadas, nos irán guiando para acometer posteriormente el estudio.

Por lo tanto, el siguiente paso será la lectura del texto original, durante la cual se va gestando un preanálisis textual, ya que existe la posibilidad de ir anotando, por ejemplo, los pasajes, el léxico, los giros que pueden plantear mayores problemas de traducción o los rasgos del estilo personal del autor, con el objetivo de comparar más adelante los elementos extraídos con las soluciones aportadas en el documento traducido.

Esta lectura revelará también las palabras clave de la obra, la relación entre escritura y lengua, el ritmo del texto y cuáles son los fragmentos significativos en los que recae el peso de la trama, habida cuenta de que una buena traducción de aquellos fragmentos será fundamental para recrear el texto original.

Recapitulando, podríamos decir que, antes de emprender el análisis de traducciones, Berman recomienda los siguientes pasos:

1. Lectura y relectura del TM:
  - a. Marcación de elementos problemáticos
2. Lectura del TO:
  - a. Preanálisis textual para identificar en el TO:
    - los elementos problemáticos de cara a la traducción
    - el estilo del autor
    - las palabras clave
    - la relación entre escritura y lengua
    - el ritmo del texto
    - los pasajes clave en la obra

Además de estos dos pasos principales, Berman alude a un tercer elemento de gran importancia: el traductor. Según él, sería interesante contar con información sobre su posición frente a la labor que realiza, las competencias y lenguas que domina, las características del proyecto de traducción en el que se enmarcan las obras estudiadas y el objetivo con el que se acomete la traslación. Así, conscientes de la importancia que suponía entrevistar al traductor de la película para poder conocer las características del proyecto, persistimos en nuestro intento de localizarlo, aunque fue en vano.

Una vez recopilados los datos, se procederá al análisis de la traducción, que debe ser entendido como un proyecto integrador, es decir, deben sumarse a él otras traducciones de la obra seleccionada, si las hubiera. Y en el supuesto de que las haya, pueden ser de tres tipos:

- 1.** Traducciones anteriores en el tiempo, pero coincidentes en lengua de destino.
- 2.** Otras traducciones contemporáneas a la misma lengua de destino.
- 3.** Traducciones a otras lenguas.

El hecho de agregar otras traducciones al estudio aumentará el valor pedagógico, ya que las "soluciones" para una cuestión específica aportadas por cada traductor son diferentes unas de otras, no solo porque cada profesional tiene una experiencia concreta, unas estrategias y una forma de expresarse, sino porque cada uno está determinado por las pautas que se recogen en su encargo de traducción. Como ya hemos mencionado, en el caso que nos ocupa, la película que se emitió en Argentina sólo difería de la de España en el título, no en la variedad empleada.

Al finalizar la acotación del *corpus* con el que se va a trabajar, pasaríamos a la confrontación: las expresiones y los pasajes o términos seleccionados en el TO, que creemos que pueden plantear dificultades a la hora de traducir, se comparan con las soluciones que figuran en el/los TM; los fragmentos, el léxico o los giros que llaman nuestra atención al leer el TM por falta de naturalidad o por cualquier otro motivo, se comparan con sus equivalentes en el TO, permitiendo así la confrontación bidireccional, es decir, se compara el TO con el TM y viceversa; se cotejan otras traducciones de la misma obra y, por último, se revisa la traducción de acuerdo con el proyecto en el que se engloba, el que, en última instancia, estará relacionado con las subjetividades que aporte el traductor, pues no hay que olvidar que, si se plantearan dos proyectos casi idénticos a dos profesionales de la traducción, el resultado propuesto por cada uno de ellos sería diferente.

En líneas generales, Berman propone esta metodología que nosotros hemos hecho nuestra, adaptando las etapas planteadas a las necesidades del archivo textual que tenemos y a las condiciones reales con que contamos. Esto es:

- Visualizar el filme doblado al español en primer lugar, como si de un original se tratara.
- Anotar las expresiones, los términos o los fragmentos "problemáticos" que nos llamen la atención.

- Visualizar el filme original.
- Anotar peculiaridades de los diálogos, expresiones que puedan plantear cierta dificultad, características del habla, etc. Intentar captar el “estilo del autor”, tal y como aconseja Berman, que en este caso, sería del director y de los personajes.
- Trabajar sin la información que podría habernos facilitado el traductor sobre el encargo y las decisiones tomadas, ya que no supimos quién realizó la traducción.
- Trabajar sin otras versiones al español simplemente porque no existen.

#### **4. HERRAMIENTAS, CONVERSIÓN DE FORMATOS Y ADAPTACIÓN**

Con el fin de adecuar el archivo textual a las distintas herramientas que utilizamos en este trabajo, hemos realizado un largo proceso de limpieza de textos, conversión y ajustes de formato que detallamos a continuación.

En primer lugar, recordemos que la extensión con que recibimos de la productora Pathé el guion de preproducción de *Bienvenue chez les Ch'tis* era .fdr (**BienvenuechezlesCh'tisV3du13-4-2007.fdr**). Este tipo de archivo, como hemos mencionado anteriormente, es propio de la herramienta de creación de guiones Final Draft y, por tanto, es necesario adquirir la licencia de este *software* privativo para poder abrir el documento. En nuestro caso, visitamos el sitio web de la empresa y descargamos la versión 8 demo del programa que han puesto a disposición del usuario con numerosas funciones y pocas restricciones, aunque algunas, como veremos a continuación, pueden llegar a ser muy importantes.

Para nosotros, la secuencia de trabajo comienza con la modificación de los diálogos recogidos en el guion de preproducción para adaptarlos a los de postproducción. Una vez hemos comprobado que el guion final se corresponde con los diálogos de pantalla, intentamos descargar la nueva versión del archivo. Y aquí nos enfrentamos a la primera restricción: si el documento cuenta con más de 15 páginas (el guion de *Bienvenue chez les Ch'tis* lo componen 112 páginas), la versión demo no permite guardar los cambios en un archivo .fdr. La única solución es guardar el documento, con las modificaciones pertinentes hechas, como un archivo .pdf. Lo hacemos y lo denominamos **BienvenuechezlesCh'tisV3du13-4-2007modificado.pdf**. En este momento, nos encontramos con el segundo problema: al ser una versión demo, aparece una marca de agua tres veces por página que reza:

**Final Draft 8 Demo**

En principio esto no parece un contratiempo, pues los diálogos se leen perfectamente, pero pronto nos damos cuenta de que, al intentar convertir el documento en un archivo de Word (.doc) o de texto plano (.txt) para hacerlo compatible con nuestras herramientas, el formato se desconfigura por la marca de agua. Además, si decidiéramos mantener la marca de agua a la hora de subir los documentos a las herramientas, las palabras que la integran contarían como parte del texto que conforma el diálogo de postproducción, lo cual alteraría los datos de nuestra investigación.

Como hemos indicado anteriormente, hemos generado un archivo en .pdf denominado **BienvenuechezlesCh'tisV3du13-4-2007modificado.pdf** que contiene el diálogo real de pantalla. El objetivo es convertir este documento en un archivo de Word y limpiarlo de todas las marcas residuales que en él pudieran permanecer como consecuencia del proceso y, por supuesto, eliminar la marca de agua proveniente del programa Final Draft 8 Demo. Para ello, utilizamos la versión profesional de Adobe Acrobat y convertimos el archivo .pdf en .doc.

Una vez convertido, procedemos a eliminar cualquier marca o etiqueta que pueda ocasionar problemas a la hora de trabajar con el archivo y las herramientas AntConc y WinAlign. En concreto, suprimimos:

- Marca de agua: Final Draft 8 Demo.
- Encabezado: repetición del título de la película en cada una de las páginas del guion.
- Pie de página: la secuencia para el pie de página consiste en el número de la página seguido de *Pathé Renn - Hirsch - Ch'timiVersion III du 13 04 07*.
- Numeración de las escenas.
- Espacios o saltos de página innecesarios.
- Toda la información existente en el guion referida a localizaciones de escenas o pautas para los personajes con intención de orientar su actuación en esa determinada escena. En cualquier caso, estos datos son superfluos para nuestro trabajo. He aquí algunos ejemplos de elementos innecesarios del guion que hemos decidido eliminar:

*1. EXT/JOUR. RUE CASSIS 1*

*Petite rue de Cassis. Un gamin, RAPHAEL, 7 ans, déboule gaiement en sautant d'un trottoir à l'autre. On découvre la mère de l'enfant, JULIE, qui a l'air dépressive et angoissée.*

*9A. EXT.JOUR MAISON GARRIGUE VIEUX MARSEILLAIS 9A*

*PHILIPPE arrive en voiture près d'une maison perdue dans la garrigue.*

*10. INT.JOUR MAISON GARRIGUE VIEUX MARSEILLAIS 10  
PHILIPPE est invité par une vieille femme à entrer dans une pièce  
sombre. Ambiance Marlon Brando dans "ApocalypseNow".*

Terminada esta fase y habiendo obtenido el guion de postproducción en francés en formato .doc y limpio, procedemos a denominar el archivo como **BienvenuechezlesCh'tisV3du13-4-2007modificado-limpio.doc**.

A partir de este momento, podemos trabajar de forma paralela con los archivos en francés y en español. Recordemos que los diálogos de las versiones doblada y subtitulada al español fueron transcritos por nosotros, por lo que directamente se ajustan a los diálogos de pantalla, están limpios y en formato .doc.

Una vez hemos conseguido un archivo textual uniforme, es decir, los tres guiones en formato .doc, limpios y fieles a los diálogos de pantalla, procedemos a trabajar con los programas elegidos.

La segunda de las herramientas que vamos a utilizar es AntConc, gestor de *corpus* lingüístico multiplataforma y gratuito. En este caso, solo analizaremos el guion en francés, pues nos interesa extraer el léxico relacionado con el dialecto *ch'timi* y con referencias culturales. La extensión compatible con AntConc es .txt, es decir, un archivo de texto plano. Teniendo en cuenta que en este momento ya contamos con todo nuestro archivo textual en formato .doc, no plantea ningún problema hacer la conversión del documento que nos interesa a .txt. No hay más que abrir el archivo en cuestión y seguir la secuencia:

Archivo > Guardar como > Guardar como tipo > Texto sin formato (\*.txt) > Guardar

Aunque nos aparece una advertencia, que nos indica que se perderá el formato, las imágenes y los objetos, pulsamos el botón Aceptar.

En el instante en que convertimos el guion en francés de .doc a formato .txt (**BienvenuechezlesCh'tisV3du13-4-2007modificado-limpio.txt**), ya está listo para ser analizado con AntConc.

El tercer programa al que vamos a subir nuestro archivo textual es WinAlign, herramienta de alineación integrada en el paquete de SDL Trados. El objetivo es alinear los segmentos del guion en lengua francesa que nos interesan para esta investigación con sus equivalentes en las versiones doblada y subtitulada en español.

El formato que hemos elegido para trabajar con la herramienta de alineación es .rtf. Como ya disponemos de todo nuestro material en .doc, la conversión a archivo de texto enriquecido (.rtf) es muy sencilla. Los pasos que hay que seguir son exactamente iguales que los realizados para la conversión a .txt, pero, en vez de seleccionar esta última extensión, debemos elegir la que ahora nos interesa:

Archivo > Guardar como > Guardar como tipo > Formato RTF (\*.rtf) > Guardar

En una primera fase, cargamos el guion en francés y el de la versión doblada en español y alineamos los segmentos seleccionados del guion en francés con sus equivalentes. En una segunda fase, subimos de nuevo el guion en francés y el de la versión subtitulada en español y realizamos idéntica acción.

Conviene subrayar que no siempre hay que respetar la alineación automática que realiza el programa, habida cuenta de que esta atiende principalmente a parámetros como la puntuación. A veces, es muy necesario ampliar o reducir los segmentos que WinAlign nos ha propuesto, con el fin de que abarquen contexto pertinente para la comprensión del léxico. En las ocasiones en que se tome esta decisión, la ampliación o reducción de los segmentos se debe llevar a cabo en ambas lenguas.

Durante todo el proceso de alineación, tenemos que mantener muy presente el objetivo de esta tarea, esto es contextualizar el léxico extraído con AntConc, enlazarlo con su equivalente en español y obtener unidades de traducción compuestas por segmentos en francés y en español, y no realizar una alineación estándar para alimentar una memoria de traducción, por lo que nos podemos tomar las licencias de expansión y reducción de segmentos mencionadas.

Un ejemplo de lo que acabamos de exponer sería el que a continuación mostramos. Si se tratara de un proyecto de alineación creado para alimentar una memoria de traducción, seguramente la oración en francés y su equivalente en español no constituirían una única unidad de traducción sino que se trataría de dos, ya que el segmento en francés se compone de dos oraciones:

*<TrU>*

*<Quality>100*

*<CrU>ALIGN!*

*<CrD>22012013, 13:44*

<Seg L=FR-FR>J'vous ai reconnu à vot' plaque qu'est 13 pas 59. Je vous ai fait chigne pour arrêter vot' carète, mais vous ne m'avez pas vu.

<Seg L=ES-EM>Reconocí su matrícula que pone 13 y aquíesh 59, le hice señas para que parara el carro pero no me vio.

</TrU>

Como el objetivo es la creación de unidades de traducción que nos permitan comparar TO y TM, es más interesante crear segmentos más amplios en ambas lenguas que recojan el contexto que nos interesa.

## 5. CONCLUSIONES

La Lingüística de Corpus y las nuevas tecnologías van de la mano en la investigación en Traducción Audiovisual, no pudiendo concebirse la una sin las otras. Por un lado, la primera aporta la base necesaria que da solidez a la configuración del *corpus* o del archivo textual; por otro, las herramientas presentadas brindan la oportunidad de extraer el máximo jugo de los archivos con los que se trabaja, gracias a la disección tanto del TO como de los TM y a la creación de unidades de traducción que permiten analizar en profundidad y en contexto las unidades léxicas en los segmentos seleccionados.

Además, la metodología que Berman (1995) propone sirve como esqueleto para el análisis de traducciones tanto de obras literarias como de material audiovisual, ya que las pautas, al ser generales, se pueden adaptar al formato en el que se encuentra el contenido objeto de estudio.

En resumen, teoría, metodología y nuevas tecnologías forman la combinación ideal para una investigación sistemática y bien organizada en Traducción Audiovisual.

## BIBLIOGRAFIA

Anthony, L. (2005). AntConc: Design and Development of a Freeware Corpus Analysis. Toolkit for the Technical Writing Classroom. En Institute of Electrical and Electronics Engineers (IEEE) (Eds.), *2005 IEEE International Professional Communication Conference: (IPCC): Limerick, Ireland, July 10-13, 2005*. (pp. 729-737). Piscataway (Estados Unidos): IEEE. Consultado el 29 de mayo de 2018 de:

<<http://www.uibk.ac.at/tuxtrans/docs/antconc.pdf>>.

Berman, A. (1995). *Pour une critique des traductions: John Donne*. París: Gallimard.

Cameo Media (Prod.) (2009). *Bienvenidos al Norte* [Película]. Barcelona: Cameo Media.

Chaume, F. (2001). Más allá de la lingüística textual: cohesión y coherencia en los textos audiovisuales y sus complicaciones en traducción. En Duro Moreno, M. (Coord.), *La traducción para el doblaje y la subtitulación*. (pp. 65-80). Madrid: Ediciones Cátedra.

Chaume, F. (2004). *Cine y traducción*. Madrid: Ediciones Cátedra.

Chaume, F. (2012). *Audiovisual translation: Dubbing*. Manchester (Reino Unido): St. Jerome Publishing.

Corpas Pastor, G. (2001). Compilación de un corpus *ad hoc* para la enseñanza de la traducción inversa especializada. *Trans*, 5, 155-184. Consultado el 29 de octubre de 2018 de:

<[http://www.trans.uma.es/Trans\\_5/t5\\_155-184\\_GCorpas.pdf](http://www.trans.uma.es/Trans_5/t5_155-184_GCorpas.pdf)>.

Corpas Pastor, G. (2008). *Investigar con corpus en traducción: los retos de un nuevo paradigma*. Frankfurt am Main (Alemania): Peter Lang.

Final Draft 8 (2011). Consultado el 29 de octubre de 2018 de: <<http://trial.finaldraft.com/>>.

Fuentes Luque, A. (2000). *La recepción del humor audiovisual traducido: estudio comparativo de fragmentos de las versiones doblada y subtitulada al español de la película "DuckSoup", de los hermanos Marx*. (Tesis Doctoral). Universidad de Granada, Granada. Consultado el 29 de octubre de 2018 de:

<<http://digibug.ugr.es/bitstream/10481/32330/1/TesisFuentesLuqueA.pdf>>.

Parodi, G. (2010). *Lingüística de corpus: de la teoría a la empiria*. Madrid: Iberoamericana Editorial Vervuert.

Pathé ! (Prod.) (2008). *Bienvenue chez les Ch'tis* [Película]. París: Pathé !



Real Academia Española (RAE) (2005). *Diccionario panhispánico de dudas*. Madrid: Real Academia Española. Consultado el 29 de octubre de 2018 de: <<http://www.rae.es/recursos/diccionarios/dpd>>.