

# La traducción como proceso de decisiones: su reflejo en los corpus paralelos<sup>1</sup>

*Petr Čermák*

Universidad Carolina, República Checa

[Petr.Cermak@ff.cuni.cz](mailto:Petr.Cermak@ff.cuni.cz)

## **Resumen:**

La idea de Jiří Levý de concebir la traducción como un proceso de decisiones es muy interesante tanto para los especialistas en traducción como para los lingüistas. El interés por analizar las traducciones, por comparar la expresión de un mismo contenido en varias lenguas dio lugar a los llamados corpus paralelos, un tipo de corpus lingüístico en el que cada texto tiene su versión correspondiente en otro u otros idiomas. Sirviéndose del material lingüístico extraído del corpus paralelo *InterCorp* (corpus creado en la Facultad de Letras de la Universidad Carolina de Praga, [www.korpus.cz/intercorp](http://www.korpus.cz/intercorp)), el artículo ilustra las posibilidades que los corpus paralelos ofrecen a los investigadores que pretenden describir el proceso de la traducción. Partiendo del análisis de traducciones que forman parte del corpus, el autor del artículo comenta diferentes factores que influyen en la traducción, en “el proceso de toma de decisiones” (el contexto, el papel del propio traductor, etc.). Además, presta atención al problema de la equivalencia/correspondencia del texto traducido con el texto original. El artículo muestra que, aunque esta nueva metodología presenta también algunas insuficiencias, los corpus paralelos constituyen en nuestro tiempo una aportación relevante a la teoría de la traducción.

**Palabras clave:** corpus paralelo, traducción, *InterCorp*, comparación de idiomas, lingüística contrastiva.

## **Translation as a decision-making process: its reflection in a parallel corpus**

### **Abstract:**

Jiří Levý's idea to conceive translation as a decision-making process is of great interest to both linguists and specialists in translation and. The interest in analyses of translations, in comparing expressions of the same content in various languages, led to the origin of the so-called parallel corpus, a type of linguistic corpus in which each text has a corresponding version in one or several other languages. Based on linguistic material obtained from the *InterCorp* parallel corpus (corpus created by the Faculty of Arts, Charles University in Prague, [www.korpus.cz/intercorp](http://www.korpus.cz/intercorp)), the article analyzes the possibilities offered by parallel corpora to researchers who aim to describe translation processes. The author of the study focuses on various factors which influence the translation as a decision-making process (e.g. the context, the role of a translator, etc.). Besides, he concentrates on the problem of equivalence/correspondence of the translated with the original. There are some

---

<sup>1</sup> Este artículo forma parte del proyecto “Program rozvoje vědních oblastí na Univerzitě Karlově č. P10 Lingvistika”, subprograma “Románské jazyky ve světle jazykových korpusů» (Lenguas románicas a la luz de los corpus paralelos”).

deficiencies regarding this methodology, but it presents an important contribution to the theory of translation.

**Keywords:** parallel corpus, translation, *InterCorp*, comparison of languages, contrastive linguistics.

### **La traduction comme un processus de prise de décision : exemple dans un corpus parallèle**

#### **Résumé :**

La conception de Jiří Levý de la traduction comme un processus de prise de décision représente un sujet d'intérêt autant pour les spécialistes de la traduction que pour les linguistes. Les « corpus parallèles », type de corpus linguistique où chaque texte possède un équivalent dans une ou plusieurs langues différentes, apparaissent grâce à l'intérêt pour l'analyse des traductions et pour la comparaison des expressions d'un même contenu entre plusieurs langues. En tenant compte du matériel linguistique obtenu du corpus parallèle *InterCorp* (corpus développé au sein de la Faculté de Lettres de l'Université Charles de Prague), cet article analyse les multiples usages des corpus parallèles pour les chercheurs qui visent à décrire le processus de traduction. On s'occupe des divers facteurs intervenant dans la traduction comme un processus de prise de décision tels que : le contexte et le rôle du traducteur. De plus, on attire l'attention sur le problème de l'équivalence/correspondance entre le texte traduit et le texte original. Bien que cette nouvelle méthodologie ait quelques faiblesses, les corpus parallèles constituent une contribution significative à la théorie de la traduction.

**Mots clés :** corpus parallèles, traduction, *InterCorp*, comparaison entre langues, linguistique contrastive.

## **1. Introducción**

Uno de los aspectos más importantes del paradigma científico moderno lo constituye su carácter internacional: los investigadores se dirigen al público internacional y lo hacen en una lengua comprensible para todo el mundo, mayoritariamente en inglés. Pero no siempre ha sido así. La historia de los países que utilizan una lengua con un reducido número de hablantes cuenta en ocasiones con pensadores muy interesantes, desconocidos total o parcialmente en el resto del mundo. La ciencia checa nos ofrece varios ejemplos de este tipo. Incluso el conocimiento de las ideas y de las obras nacidas en el famoso Círculo Lingüístico de Praga se ve influido por este fenómeno: si consideramos solo los trabajos escritos en inglés, francés o alemán, y se dejan de lado las fuentes en checo, nuestra visión del Círculo será necesariamente incompleta y deformada. Muchas ideas contenidas en los textos checos de Vilém Mathesius y otros miembros del Círculo son realmente interesantes si se tiene en cuenta el hecho de que nacieron en los años treinta y cuarenta del siglo XX y constituyen un material sustancial no solo para los historiadores de la lingüística, sino muchas veces también para los actuales especialistas en los temas que ellos trataron.

Lo mismo puede decirse, en nuestra opinión, de la obra de Jiří Levý, que, concebida en el contexto científico de los años cincuenta y sesenta del siglo pasado, sorprende por su

punto de vista moderno y poco usual en aquella época<sup>2</sup>. Es paradójico que sea precisamente la lengua en la que fue escrita lo que complica o imposibilita el acceso a una obra dedicada a la traducción, a la teoría de la traducción.

La publicación de este volumen es una prueba más de que la obra traductológica de Jiří Levý sigue siendo actual a pesar de los cincuenta años que han pasado desde su publicación. En su país es considerado el decano de la traductología checa (para más detalles, véase Králová, 2013a y 2013b), por lo que los traductólogos checos, en las últimas décadas, tratan de acercar sus ideas al público mundial (compárense Králová & Jettmarová, 2008; Králová & Cuenca Drouhard, 2013, y otros), encontrando eco en los especialistas extranjeros (Vega Cernuda, 1994 y 2011; Venuti, 2000; Navarro Domínguez, 2011, etc.).

Este artículo pretende estudiar las ideas de Levý desde el punto de vista de la lingüística de corpus, esto es, analizarlas partiendo de una metodología y unas herramientas que no existían en los tiempos de Levý. En concreto, aplicará un enfoque propio de los corpus paralelos o comparables, que es la rama metodológica de la lingüística de corpus relacionada con la lingüística contrastiva. De ello se desprende que nuestro estudio no quiere ser puramente traductológico, sino que pretende ofrecer una visión multidisciplinar de las ideas de Levý, conciliando procedimientos típicos de la lingüística de corpus, de la lingüística contrastiva y de la traductología. Lógicamente, se estudiarán solo las partes de la teoría de Levý que tienen que ver con la lengua y quedarán de un lado sus aspectos puramente literarios, si ello es posible: muchas veces una separación es forzosa y deforma la realidad.

La estructura del artículo es la siguiente: primero se estudiarán los tipos de corpus que pueden utilizarse para comparar idiomas y analizar la traducción (apartados 2.1., 2.2.), y especialmente el corpus paralelo *InterCorp*, promovido por la Universidad Carolina de Praga (apartado 2.3.). Después se repasarán algunas ideas concretas de Levý y se comentarán desde el punto de vista de la lingüística de corpus y de la lingüística contrastiva (apartado 3). Se prestará un interés especial a la concepción de la traducción como proceso de decisiones: utilizando ejemplos concretos del material sacado del corpus, intentaremos ilustrar este interesante concepto. Finalmente se resumirán las conclusiones más importantes del texto (apartado 4).

En el artículo utilizamos las versiones españolas de los textos de Levý, si aparecen en Králová & Cuenca Drouhard, 2013. Si no disponemos de versiones españolas, trabajamos con los textos originales checos (Levý, 1983 y 1996).

---

<sup>2</sup> Králová (2013a) dice al respecto: “[...] una concepción que puede parecer un poco anticuada, pero que no ha dejado de ser inspiradora para varias generaciones de investigadores en la materia, por lo que merece encontrar su lugar también en la historia de nuestra disciplina en ebullición” (p. 24).

## 2. La comparación de idiomas, la traducción y el corpus

El nacimiento de los corpus lingüísticos cambió por completo la metodología lingüística: hay tipos de estudios lingüísticos que no pueden pasarse sin la utilización de los corpus. Además, el permanente desarrollo de los instrumentos técnicos amplía constantemente el alcance de los corpus. No obstante, hay disciplinas lingüísticas que por su carácter solo difícilmente pueden aprovecharse de esta nueva herramienta y de la metodología que la acompaña. Hasta hace poco la lingüística contrastiva era una de ellas: los investigadores que pretendían sacar conclusiones cuantitativas relacionadas con la comparación de diferentes idiomas se veían obligados a limitar sus investigaciones a un número reducido de textos, procesable manualmente. Básicamente, se les ofrecían dos modos de proceder a la hora de comparar la naturaleza de un fenómeno lingüístico en dos o más idiomas: 1) trabajar con su propia intuición lingüística, es decir, establecer las equivalencias entre varios idiomas partiendo de su conocimiento personal de ellos; 2) crear un pequeño corpus personal, basado en textos originales y sus correspondientes traducciones. Es evidente que el primer procedimiento mencionado puede llegar a unas conclusiones relevantes y muy acertadas que den cuenta de las relaciones sistémicas entre los idiomas estudiados (tal como refrendan muchos estudios comparativos interesantísimos que vienen de la época anterior a la creación de los corpus informáticos), pero difícilmente puede resultar en un estudio cuantitativo defendible. Por otra parte, el segundo modo de proceder sí que trae datos cuantificables, pero los resultados tienen inconvenientes importantes: la cantidad de datos es insuficiente, los textos suelen constar de unos pocos idiolectos personales, etc. Además, el procesamiento de datos es muy exigente y laborioso, por lo que el riesgo de que el investigador cometa un error es relativamente grande<sup>3</sup>.

Ahora bien, dada la utilidad de los corpus, no es de extrañar que se buscaran métodos para aplicarlos también al campo de la comparación de lenguas. El problema básico fue el siguiente: ¿cómo obtener textos en diferentes idiomas que fueran procesables por las herramientas del corpus, reflejaran equivalencias descriptibles entre los idiomas y se prestaran a un análisis lingüístico que pudiera llevar a conclusiones relevantes? Nacieron dos concepciones básicas que parten de puntos de partida bien diferentes:

1) Hay que trabajar con textos en diferentes idiomas que compartan el registro, la temática, la procedencia regional, etc., pero que no sean traducciones. O sea, los textos en ambos idiomas deben ser originales, creados espontáneamente por hablantes nativos de cada idioma.

2) Hay que trabajar con textos traducidos: el corpus constará de textos originales y sus traducciones a otros idiomas. La comparabilidad de las versiones la garantizará

---

<sup>3</sup> Para evitar el malentendido tenemos que volver a subrayar que estamos utilizando la perspectiva de la lingüística contrastiva, no la de la traductología: el punto de vista de un traductólogo y su consideración de los métodos que acabamos de describir pueden ser radicalmente diferentes.

precisamente el hecho de que una versión sea traducción de la otra. Los textos del corpus no solo serán del mismo tipo (el mismo registro, la misma temática, etc.), sino que también serán (más o menos) idénticos semánticamente.

Es evidente que lo que marca la diferencia entre los dos métodos es la diferente apreciación de la traducción: para el primer método, que hizo nacer los llamados *corpus comparables*, un texto traducido difiere por su naturaleza del texto escrito como original. Es imposible, pues, comparar dos textos así, ya que son sustancialmente diferentes. El segundo método, punto de partida de los llamados *corpus paralelos*, presupone algo bien distinto. El trabajo con las traducciones no deforma los resultados de la comparación, todo lo contrario: una comparación así no solo ofrece todo lo que puede ofrecer un corpus comparable, sino, además, gracias a la equivalencia semántica de los textos se multiplica el aprovechamiento del corpus. Uno de estos efectos lo constituye la posible explotación del corpus paralelo para las investigaciones traductológicas: está de más subrayar que los corpus paralelos constituyen una fuente importante de material lingüístico para los que describen el proceso de la traducción. Por otra parte, en ningún caso puede sostenerse que, por definición, los corpus comparables carezcan de interés para un traductólogo: como se verá más adelante, hay un tipo de corpus comparable muy explotado por los especialistas en la traducción.

### 2.1. Corpus comparables

Como ya se ha dicho, los corpus comparables constan solo de textos originales. La posibilidad de comparar dos idiomas la garantiza la configuración de los corpus: sus autores primero fijan las características específicas de los textos del futuro corpus (el género de los textos, la función comunicativa, el registro, la temática), etc., y solo después seleccionan textos originales de ambos idiomas que cumplen con los requisitos predefinidos. El uso posible de un corpus así se desprende de sus características: muchas veces se explotan para crear diccionarios específicos de un campo léxico (por ejemplo, de una ciencia concreta, etc.). El carácter demasiado específico de los corpus comparables (el hecho de que se limiten solo a los textos que cumplen con las características definitorias) puede considerarse su punto débil. Hay intentos de reducir este inconveniente combinando más corpus comparables con diferentes características predefinidas y creando un corpus que suele denominarse *balanceado* (*balanced*), pero su creación constituye una tarea muy complicada desde el punto de vista teórico y práctico (para más detalles, véase Chlumská, 2014, y Čermák & Nádvorníková, 2015).

Como ya se ha dicho, el modo como nace un corpus comparable parece indicar que no se puede aprovechar para un análisis traductológico. Ello vale para los corpus comparables *multilingües*, que contienen textos en diferentes idiomas, pero no para los *monolingües*, lo que podría resultar paradójico si se toma en cuenta el término. Los corpus comparables monolingües constan de textos de una lengua, que se dividen en dos partes: textos originales y textos traducidos. Estos, lógicamente, no son

traducciones de aquellos: son dos corpus independientes que comparten las características predefinidas, tal y como suele pasar en los corpus comparables (o sea, el registro, la función comunicativa, etc.), pero que difieren en el modo de cómo nacieron (una parte contiene solo textos originales, la segunda, solo traducciones). Es evidente que –utilizando los términos de Levý– un corpus así no puede dar cuenta del transvase como proceso, pero sí que puede servir para un análisis de la traducción como resultado (véase el apartado 3). Dicho de otra manera, un corpus así ofrece un instrumento magnífico, por ejemplo, a los que quieren averiguar si la lengua de un texto original difiere de la lengua de un texto traducido<sup>4</sup>.

## 2.2. Corpus paralelos

Como ya se ha dicho, los corpus paralelos se caracterizan por estar constituidos de textos “idénticos” en dos o más idiomas, es decir, por estar formado por textos originales y sus correspondientes traducciones. La palabra “idénticos” va entre comillas: Jiří Levý aporta muchas observaciones interesantes acerca de este tema crucial de la traductología, comentando también cómo se concebía la “fidelidad” de la traducción en diferentes épocas de la historia de la traducción. En la lingüística de corpus de hoy se da por evidente que nunca puede haber una equivalencia total, semántica, estilística, etc., de las dos versiones del texto<sup>5</sup>. La objeción fundamental de los que critican el punto de partida de los corpus paralelos tiene que ver con lo que ya hemos comentado al hablar de los corpus comparables: un texto traducido tiene carácter particular, dado por el proceso de su creación. La lengua del texto original influye, aunque sea solo subconscientemente, en el traductor, por lo que la lengua traducida difiere de la lengua creada espontáneamente: el español del texto traducido y el español del texto original son dos lenguas diferentes.

## 2.3. El corpus paralelo *InterCorp*

Todos los comentarios acerca de las ideas de Levý que vienen a continuación parten del uso del corpus paralelo *InterCorp*<sup>6</sup>. Este corpus forma parte del ambicioso proyecto del Corpus Nacional Checo (*Český národní korpus*<sup>7</sup>), promovido por la Facultad de Filosofía y Letras de la Universidad Carolina de Praga, que cuenta con numerosos corpus del checo y de otros idiomas.

El corpus *InterCorp* nació en 2005. Su versión actual (versión 8<sup>8</sup>) reúne datos textuales en 38 lenguas acompañados de sus equivalentes checos y contiene en total 1.390

<sup>4</sup> Los interesados pueden consultar el corpus *Jerome* (<http://ucnk.ff.cuni.cz/jerome.php>), corpus comparable monolingüe que sirve para analizar la lengua checa.

<sup>5</sup> Para muchos investigadores es un argumento importante en contra de la metodología de los corpus paralelos.

<sup>6</sup> Ver: <http://ucnk.ff.cuni.cz/intercorp/>.

<sup>7</sup> Ver: <http://wiki.korpus.cz/doku.php/en:cnk:uvod>.

<sup>8</sup> Precisamente en el momento de la redacción de este artículo se prepara la puesta en marcha de la versión<sup>9</sup>, ampliada y extendida.

millones de palabras. La lengua pivote es el checo, lo que quiere decir que todos los textos existen en checo y paralelamente en uno o más lenguas. Se desprende de eso que existen no solo los 37 subcorpus creados por el checo y por cada lengua participante (o sea, subcorpus checo-español, checo-francés, etc.), sino también subcorpus que combinan dos o más idiomas diferentes del checo (subcorpus español-inglés, por ejemplo).

Los que quieren analizar el proceso de la traducción sirviéndose del *InterCorp* deben tomar en cuenta el hecho de que el corpus consta de dos grandes subcorpus. Por una parte, hay textos literarios, checos o de otros idiomas, y sus traducciones al checo o a otros idiomas. Por diferentes razones, este subcorpus, denominado *núcleo* y menor en tamaño que el otro, es idóneo para un análisis traductológico: en él son accesibles todos los datos relacionados con cada texto (autor del original, traductor, año de la edición del original, año de la publicación de la traducción, etc.); todos los textos han sido sometidos a un control manual para evitar errores de transcripción, etc. El otro subcorpus contiene material lingüístico que procede de bases de datos asequibles en internet. Esta parte del corpus –denominada *colecciones*– consta de textos periodísticos e informativos de las páginas *Project Syndicate*<sup>9</sup> y *VoxEurop*<sup>10</sup>, textos jurídicos de la Unión Europea de la página *Acquis Communautaire*<sup>11</sup>, actas de las sesiones del Parlamento europeo (*Corpus Europarl*)<sup>12</sup>, y subtítulos de películas de la base de datos *Open Subtitles*<sup>13</sup>. Algunas de estas bases de datos contienen solo las lenguas más extendidas en el mundo, entre ellas el español. Lógicamente, un análisis traductológico del subcorpus de *colecciones* es más limitado: hay menos información sobre los parámetros de los textos, los textos no han sido alineados manualmente, etc. A pesar de eso, si el investigador tiene en cuenta estos inconvenientes y configura adecuadamente su búsqueda, obtiene un material lingüístico que puede dar cuenta de la traducción de textos administrativos y periodísticos.

Hemos constatado que la cantidad de datos constituye uno de los aportes más importantes de los corpus. La construcción de un corpus paralelo siempre es más complicada que la de un corpus monolingüe, por lo que no es de extrañar que los corpus paralelos siempre sean menores que los monolingües. No obstante, consideramos que el tamaño de los diferentes subcorpus del *InterCorp*, reflejado en la Tabla 1 (que a modo de ejemplo presenta datos relacionados con algunos subcorpus españoles) ya es suficiente para realizar análisis cuantitativos defendibles<sup>14</sup>.

<sup>9</sup> Ver: <https://www.project-syndicate.org/>.

<sup>10</sup> Ver: <http://www.voxeurop.eu/cs>.

<sup>11</sup> Ver: <http://langtech.jrc.it/JRC-Acquis.html>.

<sup>12</sup> Ver: <http://www.statmt.org/europarl/>.

<sup>13</sup> Ver: <http://www.opensubtitles.org/es>.

<sup>14</sup> Probablemente esté de más explicar que las diferencias del tamaño de distintos subcorpus de la *Tabla 1* se debe a lo que acabamos de comentar (por ejemplo, en las “colecciones” figuran solo las lenguas con un gran número de hablantes).

corpus	en total	núcleo	colecciones
checo-español	127 661 459	20 228 029	107 433 430
español-inglés	103 921 521	3 961 412	99 960 109
español-francés	86 046 522	2 615 785	83 430 737
español-italiano	74 039 392	5 401 409	68 637 983
español-portugués	95 318 258	1 305 184	94 013 074
español-catalán	4 430 145	4 430 145	0

Tabla 1. InterCorp, versión 8, subcorpus españoles (en posiciones/*tokens*)<sup>15</sup>

La Tabla 1 contiene datos relacionados con el español, ya que nos dirigimos aquí al público hispanohablante. De lo explicado hasta ahora se desprende que la tabla podría ser muy larga, podría contener todas las combinaciones posibles de las 38 lenguas presentes en el corpus<sup>16</sup>.

Para concluir esta introducción informativa sobre el corpus *InterCorp* nos gustaría subrayar que los instrumentos del corpus le permiten al usuario configurar un subcorpus ajustado a sus objetivos concretos: puede seleccionar tipos de textos (núcleo, colecciones); puede escoger la dirección de la traducción (por ejemplo, puede trabajar solo con las traducciones checas de textos españoles o viceversa); puede trabajar solo con textos de un autor o traductor concreto, etc.

### 3. La concepción de Jiří Levý a la luz del material lingüístico extraído del corpus *InterCorp*

Acabamos de juzgar las posibilidades actuales de los corpus paralelos y comparables, así como de comentar someramente las características básicas de un corpus paralelo concreto, lo que nos permite considerar si la metodología de los corpus puede aportar algo nuevo a la concepción de la traducción de Jiří Levý. Antes de detenernos en algunos conceptos concretos, nos gustaría realizar dos observaciones generales.

- 1) Los puntos de partida de Levý como traductólogo –y también su objetivo– difieren sustancialmente de los de la lingüística de corpus: mientras que Levý pretende describir el proceso de la traducción entero, tomando en cuenta aspectos culturales, históricos, etc., los corpus paralelos quieren almacenar y ofrecer

<sup>15</sup> Se trata de posiciones, *tokens*, no de palabras – las posiciones son también números o signos de puntuación, por lo que el número de posiciones es siempre mayor que el de palabras.

<sup>16</sup> Para mayor información ver: <http://wiki.korpus.cz/doku.php/en:cnk:intercorp>.

grandes cantidades de datos y proveerlos de instrumentos de búsqueda siempre mejores. El lingüista de corpus analizará los datos, los cuantificará, los interpretará ateniéndose solo a lo que puede deducirse de ellos. Puede comentar también casos concretos, pero mayoritariamente le interesarán las tendencias observables en conjuntos enormes de datos. Levý, por su parte, siempre considera traducciones concretas y los analiza como un todo, tratando de tomar en cuenta todos los aspectos mencionados. Y aunque intenta objetivar, generalizar sus observaciones parciales, creando una teoría, una concepción (que será objeto de nuestros comentarios), nunca deja de tomar en cuenta el hecho de que cada traducción es única, es resultado único del trabajo de un traductor concreto en una situación concreta. En este sentido, la concentración del lingüista de corpus en la lengua traducida, en los aspectos lingüísticos del texto traducido no es suficiente para un traductólogo, que estudia también aspectos estéticos, históricos, etc., del problema. El hecho de que en este artículo centremos nuestra atención exclusivamente a los asuntos lingüísticos necesariamente hace parcial nuestra comparación de la metodología de los corpus paralelos con una concepción traductológica.

- 2) Se desprende de ello que, aunque tanto Levý como la lingüística de corpus utilizan el concepto de *traducción*, en realidad no es lo mismo. El lingüista se refiere a la lengua traducida, al texto lingüístico creado por traducción. Un traductólogo –y Jiří Levý también– tiene una visión de la traducción y del proceso de trasvase mucho más compleja; además de aspectos lingüísticos, siempre toma en cuenta aspectos culturales y otros. Si queremos comparar los dos enfoques nos vemos obligados a extraer algunos elementos del enfoque teórico de Levý y tratarlos como temas lingüísticos (puede que haya quienes opinen que no se puede hacer una extracción así). Siempre seremos conscientes –aunque no lo mencionemos explícitamente– de que muchos de estos elementos tienen una dimensión no lingüística igualmente –o incluso más– importante.

### 3.1. La traducción como proceso y como resultado

Uno de los puntos de partida del enfoque de Jiří Levý, que –sin duda alguna– se debe a su orientación funcionalista, consiste en la diferenciación entre el trasvase como proceso y la traducción como resultado. Esta dicotomía es sustancial para la traductología, ya que permite conciliar la interpretación de la traducción como una reproducción artística y la concepción del trasvase como una actividad creadora, original<sup>17</sup>.

---

<sup>17</sup> “[...] arte de traducir como actividad creadora, encaminada a producir una obra literaria de la cultura receptora, y una actividad reproductiva y reproducible, cuyo objetivo es transmitir el original” (Krállová, 2013b, p. 32).

Por definición, un corpus paralelo no puede dar cuenta de la traducción como un conjunto: por diferentes razones, le permite ver al lector solo un párrafo de las dos versiones del texto. Resulta imposible, pues, servirnos del corpus si pretendemos caracterizar la traducción como un todo, lo que es la interpretación básica del concepto de la “traducción como resultado” observable en Levý. En realidad, su ambición es aún mayor:

[...] Levý reivindica el estudio del trasvase no solo por medio de la comparación del texto de partida con el traducido, sino que insiste en la importancia de comparar los conjuntos literarios y translativos en su totalidad, haciendo hincapié en explicar el papel que desempeñan las traducciones en la creación de la literatura universal (Králová, 2013b, pp. 31-32).

Si nos limitamos a esta interpretación básica y pretendemos analizar diversas partes del texto traducido, el corpus paralelo se convierte en un instrumento muy útil: nos ofrece la posibilidad de comparar párrafos del texto traducido con los del texto original. Este es el modo más frecuente como se explota el corpus: se comparan estructuras de ambos idiomas en el nivel de oraciones o párrafos.

El estudio del trasvase como proceso es más complicado en el corpus paralelo, pero no imposible, como intentaremos demostrar en el apartado 3.4., dedicado a la traducción como proceso de decisiones.

### **3.2. La fidelidad y la libertad en la traducción**

Como es normal en la traductología moderna, Levý casi siempre trata de una manera paralela la fidelidad y la libertad en la traducción: para él son las dos caras de un mismo problema, dos aspectos que no se pueden separar. Para la metodología de los corpus paralelos el concepto de identidad de dos versiones lingüísticas de un texto también es sustancial, pero es evidente que, aunque los términos son similares, los enfoques son diferentes.

En los corpus paralelos, la identidad semántica de las versiones que los constituyen representa su característica definitoria, su razón de ser. No obstante, los estudios que ponen a prueba la metodología de los corpus paralelos analizando fenómenos lingüísticos concretos nos hacen ver que hay que distinguir entre los niveles del texto:

- 1) al nivel del texto o del enunciado podemos hablar de la identidad de lo comunicado; el significado conceptual, cognitivo<sup>18</sup> de lo expresado en la lengua de origen y en la lengua meta es el mismo;
- 2) en los niveles inferiores (sintagmas, palabras) la equivalencia de las partes del enunciado no se puede defender, como bien refrendará cualquiera que haya traducido un texto alguna vez.

<sup>18</sup> En las demás capas del significado, la situación es más complicada.

En definitiva, los datos han obligado a los lingüistas a sustituir el término *equivalente* por el de *correspondiente*: el elemento del texto en la lengua meta corresponde al de la lengua de origen, pero no siempre equivale a este, no siempre dice lo mismo (repetimos que nos referimos a la traducción tal y como la concibe la lingüística de corpus). Como se verá más adelante cualquiera que sea el término utilizado, el concepto no se presta a un análisis fácil.

Podría decirse que para Levý los análisis parciales propios de la lingüística de corpus constituyen solo una parte del problema: lo que le interesa es la fidelidad de la traducción del texto entero, casi siempre de un texto literario. Resulta que Levý se refiere a la traducción literaria y analiza cómo los traductores concretos aprovechan su libertad de tomar decisiones relacionadas con la interpretación estilística, histórica, cultural, estética, etc., del texto original. La lingüística de corpus puede aportar muy poco en esta dirección.

Para la lingüística de corpus, la equivalencia o la correspondencia de las dos versiones se relaciona con un tema que ya hemos comentado en el apartado 2 y que hizo nacer los corpus comparables: ¿es el español (o cualquier otro idioma) del texto traducido idéntico al español del texto original, creado espontáneamente por un hablante nativo? Es evidente que para ellos es un problema sustancial, porque si la lengua “traducida” fuera diferente de la lengua espontánea, constituiría una variante de lengua especial y habría que tratarla así. Para Levý y los traductólogos este problema es de menor importancia: en la historia de las traducciones literarias hay ejemplos que nos hacen ver que la lengua de la traducción puede ser diferente y esta característica suya constituye un valor positivo de la traducción. Resulta que en este tema el enfoque de la lingüística de corpus y el de la traductología son diferentes. En los apartados que siguen vamos a comentar el de la metodología de los corpus.

Ahora bien, como ya hemos señalado más arriba, las dudas al respecto de la naturaleza de la lengua traducida se han convertido en el argumento más importante en contra de la metodología de los corpus paralelos. Hay que subrayar que el estado actual de la disciplina ya permite considerar el problema partiendo del análisis de una cantidad de datos enormes: los creadores de los corpus paralelos evitan la posible influencia del idiolecto del traductor sobre los datos del mismo modo que lo hacen los corpus monolingües: siguen el lema *more data, better data* ampliando el corpus al máximo e incluyendo textos que provienen de diferentes autores y diferentes traductores.

Los estudios que se han hecho hasta ahora partiendo del corpus *InterCorp* nos hacen ver que hay tres tipos de fenómenos que podrían interpretarse como una influencia del original sobre la lengua meta:

- 1) Errores o lapsus evidentes: por ejemplo, el traductor confunde dos palabras parecidas o relacionadas. No es de extrañar que en una cantidad de datos enorme

aparezcan unos cuantos ejemplos de este tipo. En Čermák (2010) analizamos, entre otras cosas, las traducciones checas de la construcción *al anochecer*. Apareció también la traducción *za svítání* ‘al amanecer’. Este tipo de interferencia es el único que permite aplicar un enfoque prescriptivo y que, por tanto, puede ser considerada como errónea.

- 2) Construcciones que parecen indicar la influencia del texto original: suele tratarse de “interferencias” léxicas, sintácticas y de otro tipo. Entrecorramos la palabra *interferencia* porque la valoración del fenómeno es subjetiva (más abajo hablaremos de la búsqueda de instrumentos que permitan hacerla más objetiva). A pesar de eso, muchas veces se considera como una solución traslativa “indebida”: suele suponerse que el traductor se deja llevar por el texto original y en vez de buscar soluciones propias de la lengua traducida, opta por una variante que imita la lengua original. Pero la realidad suele ser muy complicada y los datos del corpus no se someten a una descripción fácil, ya que muchos factores entran en el juego (el registro; el idiolecto del traductor, que puede resultar marcado para el lector; la concepción del traductor, que puede optar intencionalmente por la “interferencia”, etc.) Mencionemos un ejemplo: En Černíková (2015) se analizan las traducciones del gerundio español. El checo no dispone de gerundio, solo existe el transgresivo, pero su utilización actualmente es muy limitada (a muchos hablantes de hoy les suena arcaico y ni siquiera son capaces de crear correctamente sus formas). En los datos de Černíková hay en total 2971 usos del gerundio español, de los cuales 81 tienen como correspondiente checo el transgresivo. Sería simplificado atribuir estas traducciones a la interferencia indebida del gerundio español: como explica la autora, la explicación es más complicada (hay formas petrificadas del transgresivo, que siguen utilizándose; el uso del transgresivo se une con unas estructuras específicas; el traductor deliberadamente hace su traducción más “literaria”, más “anticuada”, etc.). De ello se desprende que hay que proceder con precaución a la hora de analizar los datos de este tipo. Además, tales fenómenos pueden ser evidentes (lo que suele suceder en el plano léxico) pero también bastante sutiles y difícilmente apreciables (interferencias sintácticas, pragmáticas). Probablemente hay casos de interferencias indebidas, pero la experiencia con los datos del corpus nos invita a aplicar con cautela las valoraciones prescriptivas (especialmente en el componente “núcleo” del *InterCorp*, compuesto por traducciones de textos literarios).
- 3) Según algunos autores, que tienen un criterio muy estricto (y muy problemático a los ojos de otros), las técnicas utilizadas por los traductores y analizadas en la traductología moderna (nivelación, explicitación, etc.) hacen que la lengua traducida posea rasgos que no tiene la lengua espontánea<sup>19</sup>. Hasta ahora no hay datos suficientes para confirmar o desmentir esta hipótesis: es necesario contar

<sup>19</sup> Baker (1996) habla de cuatro tipos de diferencias entre las lenguas traducida y original: la lengua traducida tiende a simplificación, explicitación, normalización y nivelación.

con datos precisos y no dejarse llevar por impresiones subjetivas, por más adecuadas que parezcan.

Ahora bien, si queremos abandonar las valoraciones subjetivas y considerar la naturaleza de la lengua traducida partiendo del análisis de los datos cuantificados, la lingüística de corpus nos ofrece herramientas muy útiles. Como ya hemos explicado en el apartado 2.1., son los corpus comparables monolingües los que podrían aportarnos los datos necesarios. En un estudio dedicado a las construcciones causativas romances *faire/fare/fazer/hacer + infinitivo* y sus correspondientes checos (Čermák & Nádvorníková, 2015, p. 23ss.) pusimos a prueba el corpus comparable *Jerome*<sup>20</sup>. Se trataba de una prueba de poco alcance, pero que puede servir como modelo para unos análisis más extensos en el futuro. En aquella ocasión nos interesamos por la frecuencia de la construcción causativa checa *nechat* [“hacer”, “dejar”, “ordenar”] + *infinitivo*: comparamos su frecuencia en dos componentes del corpus *Jerome*, una compuesta de textos originales checos y otra que consta de textos checos traducidos de otra lengua. Sometimos los resultados a un análisis estadístico que permite considerar el llamado *effect size* de la diferencia cuantitativa de las frecuencias obtenidas, es decir, que permite decidir si la diferencia de frecuencias es relevante o no<sup>21</sup>. Los datos nos hicieron ver que en este caso concreto no es posible demostrar una diferencia relevante entre el checo “original” y el “traducido”. Somos conscientes, no obstante, de que es solo un fenómeno lingüístico, además parcial y de menor importancia, por lo que no es posible sacar conclusiones contundentes. A pesar de ello, la explotación de los corpus comparables parece prometedora para los que quieren analizar rasgos particulares de la lengua traducida.

Así que es muy probable que aún tardemos mucho en tener datos suficientes respecto al carácter particular de la lengua traducida. Por eso no hay que olvidar que los corpus paralelos actuales ofrecen posibilidades que permiten hacer que nuestro análisis sea lo más objetivo posible. Mencionemos al menos tres de ellas:

- 1) En el corpus *InterCorp*, la posibilidad de seleccionar el subcorpus (“núcleo”, “colecciones”) le permite al investigador elegir un tipo de traducción. En el “núcleo”, que consta de textos literarios, la concepción traductológica del traductor puede influir mucho en la lengua meta. Por otra parte, en los textos administrativos de las “colecciones”, el objetivo del traductor es diferente, lo que se reflejará también en la lengua.
- 2) Los corpus paralelos de calidad disponen de la posibilidad de seleccionar la dirección de la traducción. Un análisis correcto de los datos debería separar las dos direcciones, es decir, tratar por separado las traducciones checas de textos españoles y las traducciones españolas de textos checos, por ejemplo. Si

<sup>20</sup> Ver: <http://ucnk.ff.cuni.cz/jerome.php>.

<sup>21</sup> La descripción del análisis, que se sirve del llamado coeficiente de Dice, aparece en Čermák & Nádvorníková, 2015, p. 24.

procedemos así tenemos la posibilidad de controlar los datos, de reinterpretarlos con los datos extraídos del subcorpus compuesto de textos traducidos en la dirección contraria. El corpus *InterCorp* dispone de instrumentos que permiten una selección muy detallada del tipo de traducción.

- 3) Los corpus paralelos y comparables están en auge, pero siguen siendo pequeños si aplicamos los criterios utilizados para los corpus monolingües. No es de extrañar pues que otro procedimiento que suele utilizarse para disminuir la posible influencia negativa de la traducción sobre los datos sea el aprovechamiento de los enormes corpus monolingües: su ingente cantidad de datos puede ayudar a rectificar conclusiones erróneas extraídas de los corpus paralelos (véase Altenberg & Granger, 2002, p. 9).

### 3.3. La importancia de la figura del traductor

La figura del traductor constituye otro tema predilecto de Jiří Levý. Como siempre, su enfoque es muy complejo y se basa en un punto de partida multidisciplinar. La lingüística de corpus puede aportar datos estadísticos sobre las traducciones de un traductor concreto y compararlos con los de otros traductores. No obstante, no es su objetivo primordial valorar las cualidades de los traductores, ni mucho menos. Lo que resulta más importante son los instrumentos que ofrecen los corpus paralelos a los investigadores que pretenden estudiar la importancia de la figura del traductor.

- 1) Los autores de los corpus paralelos consideran al traductor como un parámetro importante del corpus: no solo que lo incluyen como un dato sustancial en la ficha de cada texto, sino que también intentan contar con el mayor número posible de traductores diferentes (sabemos muy bien que la traductología tradicional se veía obligada, por razones técnicas, a trabajar con pocos datos provenientes de pocos traductores). Claro está que, en las literaturas de las lenguas con pocos hablantes, a veces se debe afrontar el problema del escaso número de traducciones: por ejemplo, el corpus paralelo español-checo es mucho más grande que el checo-español, lo que se debe a una riquísima tradición checa de traducir libros en español que contrasta con un número relativamente pequeño de traducciones españolas de libros checos. El número de libros traducidos está relacionado, lógicamente, con el número de traductores, es decir, con el número de idiolectos (y de concepciones traductológicas) diferentes.
- 2) Por el momento, el corpus *InterCorp* dispone de una única traducción de cada libro. En el futuro probablemente serán accesibles varias traducciones de un mismo libro (si las hay). Será un instrumento magnífico para los que quieran estudiar los diferentes modos de cómo se ha traducido el mismo texto (lógicamente, no será grande este subcorpus específico, contendrá solo las obras consagradas, que suelen traducirse repetidamente).

- 3) Por último, hay que señalar que el *InterCorp* le permite al usuario seleccionar traductores concretos y combinar este criterio con otros (es posible, por ejemplo, estudiar traducciones de un autor concreto, hechas por un traductor concreto).

### 3.4. La traducción como proceso de decisiones

Jiří Levý no es el único ni el primero que concibe la traducción como proceso de decisiones (en sus textos, remite a otros autores que también tratan el tema), pero su concepción –por su complejidad y por la aportación de numerosos ejemplos concretos– representaba en su época un aporte metodológico muy interesante e importante. Recordemos, en primer lugar, cuál es su punto de partida:

Desde el punto de vista teleológico, es decir, estudiada desde el punto de vista de su finalidad, la traducción es un proceso de comunicación: el objetivo de la traducción es comunicar. Desde el punto de vista del trabajo del traductor, es decir, desde el punto de vista pragmático, el trasvase es un continuo proceso de decisiones; se trata de una serie de determinado número de situaciones consecutivas, a modo de movimientos en el juego, en las cuales el traductor debe decidir entre un número de alternativas, generalmente bastante bien definidas (Levý, 1971, versión española en Králová & Cuenca Drouhard, 2013, p. 121).

En este proceso de decisiones intervienen diversos factores jerarquizados (nuestra explicación se ve inspirada por Levý, 1967 y 1971 –en su versión española publicada en Králová & Cuenca Drouhard, 2013– y en Králová, 2013b, p. 33-34):

#### 1) Instrucciones delimitativas

Son las instrucciones que delimitan el llamado paradigma, o sea, alternativas sistémicas que se le ofrecen al traductor. Por ejemplo, traduciendo la palabra inglesa *boy*, el traductor elige entre las expresiones españolas *chico*, *muchacho*, *joven*, *chaval*..., que forman el paradigma, conjunto de elementos que responden a cierta instrucción, conjunto organizado conforme a diferentes criterios (matices semánticos, niveles estilísticos, etc., ya que no son elementos equivalentes).

#### 2) Instrucciones selectivas

Son las instrucciones que el traductor aprovecha para elegir una solución del conjunto que acabamos de definir: la elección del recurso respectivo se realiza conforme a todo un conjunto de instrucciones selectivas, que son de diferentes tipos:

A) El contexto como instrucción objetiva.

B) La estructura de la memoria del traductor como instrucción subjetiva; su aplicación permite activar los paradigmas subjetivos del traductor y la

introducción de este concepto permite explicar la tendencia a la nivelación del estilo.

C) La norma estética como instrucción intersubjetiva.

Entre la instrucción delimitativa y la selectiva existe la relación de subsunción: partiendo del conjunto de alternativas definidas por la instrucción delimitativa, la instrucción selectiva aparta un subconjunto, para el cual se convierte en la instrucción delimitativa, etc., hasta llegar a un paradigma de un solo miembro.

Este breve resumen de la concepción de Levý está simplificado, pero nos permitirá comentar el tema desde el punto de vista de la lingüística de corpus. Para empezar, permítasenos presentar otra cita más de la obra de Levý:

Comprender el trasvase como proceso de decisiones encuentra su justificación en una simple razón: corresponde al empirismo, a las experiencias prácticas. Al considerarlo como proceso de decisiones, es posible describirlo exactamente aplicando el aparato formal de la teoría de los juegos y de las decisiones estadísticas. Esta formalización no aportará resultados, servirá más bien para llamar atención sobre algunos conceptos noéticos planteados por la concepción mencionada (Levý 1971, versión española en Králová & Cuenca Drouhard 2013, p. 123).

Se desprende de la cita que Levý era consciente de la posibilidad de “formalizar” la descripción del proceso de decisiones, de aplicar una metodología más exacta. No es tal vez demasiado atrevida la hipótesis de que, estar aún entre nosotros, Levý sería aficionado a la explotación de los corpus en la investigación del proceso de la traducción.

En general, vale lo que ya hemos dicho: el corpus puede suministrar datos que den cuenta de los diferentes aspectos del proceso de decisiones y puede hacerlo de una forma bastante sofisticada. Teóricamente, dos tipos de estudios son imaginables:

- 1) Análisis complejos de pasajes de textos: El investigador someterá a un análisis traductológico detallado oraciones o incluso párrafos enteros, prestando atención a todas sus partes y todos sus aspectos. Este tipo de estudios no ha sido muy frecuente hasta ahora, lo que puede deberse al hecho de que el corpus permite ver solo pasajes restringidos del texto, por lo que un contexto más extenso no es accesible. Probablemente, los instrumentos del corpus que permiten la cuantificación no se explotarán mucho.
- 2) Análisis de fenómenos parciales concretos: este tipo de estudios prevalece claramente en la explotación de los corpus paralelos y comparables. Si el investigador delimita convenientemente su objetivo y configura los instrumentos de búsqueda adecuados, el corpus le permitirá ofrecer un cuadro complejo y cuantificado del fenómeno en cuestión. En el texto que sigue nos gustaría presentar unos cuantos ejemplos de estudios de este tipo que pueden mostrar cómo se refleja el proceso de decisiones en un corpus paralelo.

Antes que nada, hay que subrayar que las instrucciones mencionadas por Levý no se reflejan de la misma manera. El corpus paralelo nos da mucha información sobre el paradigma como resultado de la aplicación de las instrucciones delimitativas (aunque, lógicamente, no nos hará ver todos los miembros del conjunto que respondan a la instrucción; por otra parte, el corpus paralelo muchas veces nos ofrece alternativas que no suelen aparecer en la descripción del fenómeno ofrecida por las gramáticas contrastivas y lo hace, además, de una manera muy cómoda para el usuario). Si el investigador quiere describir las instrucciones selectivas debe poner más ahínco, se necesita una interpretación personal del material que por definición no es fácil. La valoración de la influencia de diferentes factores es complicada, como nos hace ver el mismo Levý: de los tres factores que menciona (el contexto, la memoria del traductor, la norma estética) solo el primero se presta a un análisis directo en el corpus paralelo. Los demás son tal vez analizables de una manera indirecta y todo parece indicar que exigen un estudio translatólogo, no puramente lingüístico, por lo que no vamos a tratarlos aquí.

Veamos, en primer lugar, las instrucciones delimitativas y los paradigmas. Králová (2013b) resume la postura de Levý respecto a los paradigmas de la manera siguiente:

[...] tanto la definición de los paradigmas como la elección de los medios concretos de las alternativas respectivas resulta tanto más difícil, cuanto más elevado es el nivel de estructuración del texto delimitado: más fácil resulta en el nivel léxico (ser / estar corresponden a un solo verbo checo, *být*), después sigue el nivel morfológico (en checo y en varias lenguas eslavas, por ejemplo, existe un solo tiempo pasado), el sintáctico (en checo no existen las construcciones semipredicativas de infinitivo, gerundio y / o participio), siendo los más complicados el de estructuración del enunciado y el de los patrones textuales (p. 34).

Los estudios comparativos realizados hasta ahora mediante el corpus *InterCorp* refrendan a grandes rasgos esta afirmación de Levý: mientras que los análisis en los niveles léxico y morfológico son más fáciles, ya que la identificación del elemento correspondiente en la lengua meta no da problemas, en los niveles más altos los análisis requieren una configuración de la búsqueda más complicada. Decimos *a grandes rasgos* porque muchas veces es casi imposible la adscripción del fenómeno analizado a un nivel de lengua concreto: la solución del traductor es más compleja y afecta a diferentes niveles lingüísticos a la vez, tal como más abajo mostraremos en un ejemplo concreto.

Si aceptamos la diferenciación en niveles, buscar diferentes “equivalentes<sup>22</sup>” de una unidad léxica es fácil en el corpus *InterCorp* –el programa incluso nos ordena los equivalentes según su frecuencia. Desde hace poco el corpus dispone de un subprograma denominado *Treq (Translation Equivalents Database)*<sup>23</sup>, que es capaz de predecir automáticamente los equivalentes posibles de las palabras, combinando el checo con

<sup>22</sup> Por las razones ya explicadas entrecomillamos la palabra *equivalente*.

<sup>23</sup> <http://treq.korpus.cz/index.php>

cualquier idioma que forma parte del *InterCorp*<sup>24</sup>. Así adquirimos el paradigma de los equivalentes de la palabra buscada que refleja la presencia de la palabra en cuestión en el corpus (evidentemente, cuanto más grande es el corpus, tanto más miembros tendrá el paradigma ofrecido por el buscador). Igual de fácil es la búsqueda de paradigmas en el nivel morfológico: si nos atenemos al ejemplo de la cita de Králová (2013b, p. 34) identificar los posibles correspondientes de los tiempos pasados españoles en checo tampoco es difícil, y lo mismo vale para su búsqueda en el corpus.

Como ya hemos señalado, en los niveles más altos, la tarea de buscar el paradigma es mucho más complicada, porque los miembros del paradigma suelen ser complejos (muchas veces con una estructura interna jerarquizada), muy heterogéneos e imprevisibles. Mencionemos un ejemplo concreto. En Čermák & Nádvorníková (2015) analizamos los “equivalentes” de la construcción española *hacer + infinitivo* en checo. Aunque las gramáticas checas suelen hablar de tres o cuatro construcciones que suelen corresponderse con la perífrasis española, en los datos del corpus delimitamos ocho tipos diferentes, algunos de ellos difícilmente previsible, pero frecuentes. Mencionemos un ejemplo: el segundo equivalente checo más frecuente (21 % de todos los casos) es el tipo que refleja un cambio de estructura sintáctica (se reestructuran papeles sintácticos), un tipo poco previsible, compárese:

*¿Qué le hace pensar así? – Proč myslíte?* [literalmente: *¿Por qué piensa usted eso?*]

Resulta que en los estudios basados en el uso del corpus es muy razonable someter primero los datos a un análisis previo, identificar los miembros del paradigma (o los llamados *recurrent translation patterns*, véase Krzeszowski, 1990, p. 27) –si no procedemos así, solo difícilmente abarcaremos todos los miembros del paradigma (o por lo menos su mayoría). Dicho de otra manera: la delimitación del paradigma de Levý es un procedimiento que no solo es útil para el análisis traductológico, sino también es uno de los primeros pasos de la explotación de los corpus paralelos.

Aún nos falta comentar las instrucciones selectivas. En los estudios basados en un corpus paralelo, la búsqueda de las instrucciones selectivas constituye la fase cualitativa del análisis que sigue al análisis cuantitativo: muchas veces (aunque no siempre) el lingüista-investigador quiere explicar la distribución cuantitativa de los datos y caracterizar los factores que influyen sobre ellos. Como ya hemos dicho, de los factores mencionados por Levý, solo el contexto se presta fácilmente a un análisis basado en el corpus. El *InterCorp* nos ofrece una serie de instrumentos específicos que permiten tomar en cuenta un contexto muy amplio de la palabra buscada (suele trabajarse hasta con seis posiciones delante y detrás de la palabra pivote). Además, posee la función de identificar colocaciones, esto es, combinaciones frecuentes de palabras. En este sentido, el

---

<sup>24</sup> No es de extrañar que el programa sirva ante todo para las investigaciones lexicológicas y lexicográficas, que por definición tienen un objetivo diferente de las investigaciones traductológicas.

investigador puede apreciar la influencia del contexto sobre la selección del miembro concreto del paradigma. Como sucede siempre en los corpus paralelos, esta influencia puede cuantificarse (es decir, puede relacionarse la frecuencia de cada miembro con un contexto determinado), o puede someterse a un análisis translatólogo que tome en cuenta también otros factores importantes aparte del contexto. Igual que en las instrucciones delimitativas, también aquí el análisis es menos complicado en el nivel léxico y morfológico (siempre y cuando sea posible en el fenómeno en cuestión separar claramente los niveles lingüísticos).

Para terminar, nos gustaría, a modo de ejemplo, comentar dos estudios concretos, basados en el material lingüístico del corpus paralelo *InterCorp*, teniendo en cuenta la concepción de la traducción como proceso de decisiones de Levý. Los dos parten de una situación muy común en la comparación de idiomas: un fenómeno que está presente en un idioma no tiene su contrapartida del mismo tipo en el otro idioma, por lo que, lógicamente, el otro idioma debe servirse de recursos de otra índole.

En el pasaje de Králová (2013b, p. 34) citado más arriba se habla de la inexistencia de construcciones semipredicativas con infinitivo en checo. En Čermák (2010) analizamos los “equivalentes” de la construcción española *al + infinitivo*. Dada la ausencia del elemento del mismo tipo en checo nos vimos obligados a repasar las ocurrencias del corpus y delimitar los tipos de expresión del significado de la construcción en checo (esto es, el paradigma de Levý). En la primera fase, la tipología básica debe ser bastante general. En este caso concreto, trabajamos con cinco tipos básicos: el equivalente checo de la construcción semipredicativa tiene forma 1) de oración principal, 2) de oración subordinada, 3) de transgresivo, 4) de elemento sintáctico (por ejemplo, un sintagma compuesto por preposición + sustantivo deverbativo), 5) de una construcción diferente de las anteriores. La cuantificación de los datos nos hizo ver que el tipo 2 y 4 prevalecen claramente (50,8 % y 43,7 %, respectivamente). En la siguiente fase analizamos los factores que influían sobre la selección del tipo concreto de equivalente (la lexicalización, patente por ejemplo en las traducciones de construcciones frecuentes como *al amanecer*; el contexto, etc.), es decir, el reflejo de las instrucciones selectivas. La descripción de estos factores permitió especificar las traducciones y crear un cuadro mucho más concreto de equivalentes checos.

En otro estudio Čermák (2016) invertimos la dirección de la traducción: nos dedicamos a un fenómeno checo que no tiene correspondiente productivo español (del mismo tipo). Los llamados adjetivos procesuales constituyen un subtipo de adjetivos verbales en checo y sirven para expresar un proceso: el proceso se ha transformado en una propiedad y el verbo se ha convertido en adjetivo; por ejemplo: *muž píše dopis* “un hombre escribe una carta” → *muž písící dopis*, literalmente “un hombre <escribir-adj> una carta”. El paradigma básico de equivalentes españoles consta de seis tipos: 1) un adjetivo (*creciente, trabajador*, etc.); 2) una oración subordinada de relativo (*que se está muriendo, que colgaba*, etc.); 3) un gerundio (*una granja ardiendo*, etc.); 4) una oración principal (en el texto español se reestructuran los papeles sintácticos y el equivalente del verbo checo que engendró el

adjetivo aparece en forma personal, desempeñando el papel del predicado de una oración principal; 5) una construcción analítica (*la nieve en movimiento*, etc.); 6) no hay equivalente material del adjetivo. Los datos cuantificados nos dicen que el tipo 1 y 2 prevalecen claramente en el corpus paralelo (44 % y 28 % de todas las ocurrencias). El análisis siguiente permite especificar y precisar la tipología básica y también descubrir los factores que influyen sobre la decisión del traductor (esto es, las instrucciones selectivas): son detectables restricciones semánticas y sintácticas (por ejemplo, es muy importante la presencia de complemento o complementos del adjetivo procesual checo, ya que ello excluye muchas veces la posibilidad de un adjetivo español como equivalente).

Repitamos lo expresado más arriba: la traducción como proceso de decisiones es un concepto sumamente interesante. Todo parece indicar, no obstante, que la realización de este procedimiento teórico es muy compleja y sutil en la práctica, como demuestran nuestros estudios: los miembros de los paradigmas suelen ser numerosos y difícilmente abarcables por el corpus paralelo, por más grande que sea (lógicamente, aún menos por la mente de un traductor), lo que se debe, entre otras cosas, a su alto grado de heterogeneidad (ya en el nivel léxico pueden ser miembro del mismo conjunto denominaciones univerbales y poliverbales, y en los niveles más altos la heterogeneidad es aún más grande). Además, tal y como apunta Levý, la aplicación de las instrucciones debe ser recursiva muchas veces, debe repetirse, alternando instrucciones delimitativas y selectivas; el proceso puede ser tan complicado en la práctica, que casi no se presta a una descripción.

#### 4. Conclusiones

Aunque a primera vista el mundo traductológico de Levý y el mundo de los corpus lingüísticos pueden parecer alejados el uno del otro, esto no siempre es así. Hemos intentado analizar algunas ideas de Levý sirviéndonos de la metodología de los corpus paralelos (evidentemente, no todas se prestan a un análisis así, ni mucho menos) y hemos encontrado varios puntos comunes importantes. Mencionemos al menos dos:

- 1) La metodología de los corpus paralelos ofrece a los estudios traductológicos un novedoso material en dos o más versiones lingüísticas, así como instrumentos técnicos para procesarlo, cuantificarlo y prestar atención a toda una serie de parámetros que son importantes para un análisis traductológico.
- 2) La descripción detallada de Levý de la traducción como proceso de decisiones se basaba en una ejemplificación rica, pero solo difícilmente podía tomar en cuenta el aspecto cuantitativo, la frecuencia de los miembros de los paradigmas, etc. Los corpus paralelos pueden suministrar una cantidad de datos enorme que ponga a prueba la validez de las conclusiones de Levý sobre la aplicación de las instrucciones delimitativas y selectivas. Además, pueden simplificar la búsqueda de los miembros de los paradigmas, ya que sus datos son capaces de encontrar miembros que intuitivamente no son previsibles.

La metodología de los corpus paralelos no siempre refrenda las conclusiones de Levý. A pesar de eso, muchas ideas de este traductólogo checo son interesantes para la lingüística de corpus de hoy, por lo que podemos decir que la colaboración interdisciplinar entre la metodología traductológica de Levý y la metodología de los corpus puede ser enriquecedora para ambas partes.

## Referencias

- Altenberg, B., Granger, S. (2002). *Lexis in Contrast. Corpus-Based Approaches*. Amsterdam: Benjamins.
- Baker, M. (1996). Corpus linguistics and translation studies: Implications and applications. En Sager J. & Somers H. (eds.), *Terminology, LSP, and Translation: Studies in Language Engineering in Honour of Juan C. Sager* (pp. 175-186). Amsterdam: Benjamins.
- Čermák P. (2010). AI + infinitivo: estudio contrastivo checo-español. *Acta Universitatis Carolinae - Philologica*, 2/2009, XVIII, pp. 49-64.
- Čermák P. (2016): Acerca de una diferencia sistémica entre el checo y el español: adjetivos procesuales checos. *Romanica Olomucensia*, 28(1), pp. 23–33.
- Čermák, P. & Nádvorníková, O. (eds.) (2015). *Románské jazyky a čeština a čeština ve světle paralelních korpusů* [Las lenguas romances y el checo a la luz de los corpus paralelos]. Praga: Nakladatelství Karolinum.
- Černíková, Š. (2015): Gerundium ve španělštině [El gerundio en español]. En: Čermák, P. & Nádvorníková, O. (eds.) (2015). *Románské jazyky a čeština a čeština ve světle paralelních korpusů* (pp. 219-241). Praga: Nakladatelství Karolinum.
- Chlumská, L. (2014). Není *korpus* jako *korpus*: Korpusy v kontrastivní lingvisticke a translatoologii [No todos los corpus son iguales: los corpus en la lingüística contrastiva y en la translatoología]. *Časopis pro moderní filologii*, 96(2), pp. 221-232.
- Levý J. (1996) [1956]. *České teorie překlada 1, 2* [Teorías checas de la traducción]. Praga: Ivo Železný.
- Levý J. (1967). Translation as a Decision Process. En *To Honor Roman Jakobson II*. (pp. 1171-1182). The Hague: Mouton.
- Levý J. (1971). *Bude literární věda exaktní vědou?* [¿Será la ciencia literaria una ciencia exacta?]. Praga: Československý spisovatel.
- Levý J. (1983) [1963]. *Umění překlada* [El arte de la traducción]. Praga: Panorama.
- Králová, J. & Jettmarová, Z. et al. (2008). *Tradition versus modernity. From the classic period of the Prague School to translation studies at the beginning of the 21st century*. Praga: Univerzita Karlova v Praze y Togga.
- Králová, J. & Cuenca Drouhard, M. J. (eds.) (2013). *Jiří Levý: una concepción (re)descubierta*. Monográficos de la revista Hermēneus. Soria: Vertere
- Králová, J. (2013a). Criterios de la edición española. En Králová, J. & Cuenca Drouhard, M. J. (eds.), *Jiří Levý: una concepción (re)descubierta*. (pp. 19-24). Monográficos de la revista Hermēneus. Soria: Vertere.

- Králová, J. (2013b). Introducción a la edición española. En Králová, J. & Cuenca Drouhard, M. J. (eds.), *Jiří Levý: una concepción (re)descubierta*. (pp. 29-38). Monográficos de la revista *Hermēneus*. Soria: Vertere.
- Kreszowski, T. P. (1990): *Contrasting Languages: The Scope of Contrastive Linguistics*. Berlín: Mouton de Gruyter.
- Navarro Domínguez, F. (2011). Jiří Levý: La teoría de la traducción y la lingüística. En Králová, J. (ed.), *Posibilidades y límites de la comunicación intercultural* (pp. 21-30). Praga: Karolinum.
- Vega Cernuda, M. A. (1994). *Textos clásicos de la teoría de la traducción*. Madrid: Cátedra.
- Vega Cernuda, M. A. (2011). El caso Levý: Fenomenología de su recepción y valoración de sus aportaciones en el contexto de la traductología de la época. En Králová, J. (ed.), *Posibilidades y límites de la comunicación intercultural* (pp. 11-20). Praga: Karolinum.
- Venuti, L. (2000). *The Translation Studies Reader*. London & New York: Routledge.

### **Corpus lingüísticos:**

- Corpus *Jerome*. Instituto del Corpus Nacional Checo, Facultad de Filosofía y Letras, Universidad Carolina. Accesible en <http://ucnk.ff.cuni.cz/jerome.php>.
- Corpus paralelo *InterCorp*. Instituto del Corpus Nacional Checo, Facultad de Filosofía y Letras, Universidad Carolina. Accesible en <http://www.korpus.cz/intercorp>.