

MUTATIS
MUTANDIS

Mutatis Mutandis. Revista Latinoamericana de Traducción
ISSN: 2011-799X
Universidad de Antioquia

Tebé, Carles; Pissolato, Luciana
Reflexiones sobre la metodología terminológica en contextos de
traducción: ¿qué se pierde en el trabajo con herramientas informáticas?*

Mutatis Mutandis. Revista Latinoamericana de Traducción,
vol. 12, núm. 1, 2019, Enero-Junio, pp. 272-293
Universidad de Antioquia

DOI: <https://doi.org/10.27533/udea.mut.v12n1a11>

Disponible en: <https://www.redalyc.org/articulo.oa?id=499273335011>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

UDEM
redalyc.org

Sistema de Información Científica Redalyc
Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso
abierto

Reflexiones sobre la metodología terminológica en contextos de traducción: ¿qué se pierde en el trabajo con herramientas informáticas?

Carles Tebé

ctebe@uc.cl

Pontificia Universidad Católica de Chile, Santiago de Chile

Luciana Pissolato

lupissolato@uc.cl

Pontificia Universidad Católica de Chile, Santiago de Chile

Resumen

Los supuestos metodológicos del trabajo en terminología desde una aproximación lingüística y textualista están bien establecidos desde hace dos décadas. No puede hablarse de una metodología terminológica única, puesto que lo que denominaremos escenarios terminológicos pueden suponer recorridos y etapas metodológicas muy diferentes –por ejemplo, una recopilación de neologismos de prensa, la compilación de un glosario para la traducción, o un trabajo de normalización de términos.

La diversidad de escenarios terminológicos se corresponde con una gran variedad de herramientas disponibles para el terminólogo: gestores de bases de datos, alineadores, concordanciadores, herramientas de traducción asistida y automática, detectores de terminología, extractores de definiciones, etc. Cada escenario de trabajo exige su propio conjunto de herramientas.

En este artículo, planteamos una serie de reflexiones metodológicas en torno al trabajo en terminología: si bien las herramientas nos permiten tratar grandes volúmenes de datos, ser más consistentes, y validar hipótesis de una forma más objetivable, sobre todo cuantitativamente, ¿qué se pierde al dejar que sea la herramienta, y no el investigador, quien extraiga evidencias y presente resultados? A partir de ejemplos concretos, responderemos a las siguientes preguntas: ¿se pierden datos relevantes? ¿En qué etapas, mayormente? ¿Qué tipo de datos son y qué impacto tienen sobre la investigación? Además de datos, ¿qué otros aspectos de la metodología se pueden ver afectados? Puesto que la investigación se desarrolló en un escenario de traducción, en las conclusiones proyectaremos las consecuencias de estas reflexiones sobre la enseñanza-aprendizaje de la terminología en contextos de traducción.

Palabras clave: metodología terminológica, terminótica, lingüística de corpus aplicada a la traducción, escenarios terminológicos, encargo de traducción



Reflections on Terminological Methods in Translation Contexts: What is Lost in the Work with Computing Tools?

Abstract

The methodological principles of terminology work from a linguistic and textual point of view are well established since the last two decades. There is not a unique methodological approach in the field of terminology, since what we call in this paper terminological scenarios can lead to different methodological pathways and stages —e.g. a compilation of neologisms found in newspapers, the compilation of a glossary for translation purposes, or a standardized glossary in a technical environment.

The diversity of terminological scenarios also points to a variety of tools that are available for the terminologist: database management systems, text alignment tools, concordance programs, assisted and automatic translation tools, terminology detection systems, definition extraction systems, etc. A terminologist facing any given scenario must choose the appropriate set of tools.

In this article, we pose a series of methodological questions about terminology work: although computing tools allow the researcher to deal with big data volumes, being more consistent, and allow him/her to validate hypothesis in a more objectively way, what other important things are left behind when is the tool, and not the researcher, who extracts evidences and presents results? We will answer to the following questions: in this process, are relevant data left behind? In which stages, specifically? Which type of data are lost and what impact do they have on the research goals? Besides data, which other aspects of the methodology can be affected? Since this research is done in a translation context, in the conclusions we will project the consequences of these reflections on the teaching-learning of terminology in translation programs.

Keywords: terminological methods, terminotics, corpus linguistics applied to translation studies, terminological scenarios, translation assignment

Réflexions sur la méthodologie terminologique dans les contextes de traduction : quelles sont les pertes dans le travail avec des outils informatiques ?

Résumé

Les méthodologies en terminologie d'après une approche linguistique et textuelle sont bien établies depuis deux décennies. Nous ne pouvons pas parler d'une seule méthodologie terminologique, car ce que nous appellerons des scénarios terminologiques peut impliquer des parcours et des étapes méthodologiques très différentes, entre autres, une compilation de néologismes de la presse, la compilation d'un glossaire pour la traduction ou un travail de normalisation terminologique.

La diversité des scénarios terminologiques présuppose alors une grande variété d'outils disponibles pour le terminologue : gestionnaires de bases de données, aligneurs, concordanciers, outils de traduction assistée par ordinateur et automatique, détecteurs de terminologie, extracteurs de définitions, entre autres. Chaque scénario de travail exige son propre ensemble d'outils.

Dans cet article, nous proposons des réflexions méthodologiques sur le travail en terminologie : bien que les outils permettent de traiter de gros volumes de données, d'être plus cohérents et de valider des hypothèses d'une manière plus objective et quantitative, nous nous demandons : qu'est-ce on perd dans le travail terminologique quand c'est la machine et non le chercheur qui extrait les évidences et présente les résultats ? En prenant des cas concrets, nous répondons les questions suivantes : des données pertinentes sont-elles perdues ? À quelles étapes, principalement ? De quel type de données s'agit-il et quel impact ont-ils sur la recherche ? En plus des données, quels autres aspects de la méthodologie peuvent être affectés ? Tant que la recherche a été développée dans un scénario de traduction, nous concluons avec les conséquences de ces réflexions sur l'enseignement et l'apprentissage de la terminologie dans des contextes de traduction.

Mot-clés : méthodologie terminologique, terminotique, linguistique de corpus appliquée à la traduction, scénarios terminologiques, commission de traduction.

1. Marco teórico general y supuestos metodológicos

El marco teórico general para la redacción de este artículo es la Teoría Comunicativa de la Terminología, propuesta por Maria Teresa Cabré y su equipo de colaboradores en el Institut Universitari de Lingüística Aplicada, en la Universitat Pompeu Fabra, Barcelona, a partir de finales de los años 90, y principios de la década siguiente (Cabré, 1999, 2003; Cabré y Estopà, 2005; Freixa, 2005; Tebé, 2006; *et al.*).

La Teoría Comunicativa de la Terminología (en adelante, TCT) surge a finales de los años 90 como una alternativa al paradigma teórico de la Teoría General de la Terminología (en adelante, TGT) que en la década de los años 70 y 80 había dominado la visión de la terminología como disciplina.

Así, establece los siguientes principios y fundamentos (Cabré, 1999, 2003):

- La terminología, y, por ende, su objeto de estudio (las unidades terminológicas) forman parte del lenguaje natural. Sin embargo, no se rechaza que ciertos conceptos especializados puedan constituirse como unidades del lenguaje artificial.
- Las unidades terminológicas poseen un carácter interdisciplinario y multidimensional, al igual que la terminología. Por ende, pueden ser explicadas a partir de las tres dimensiones: lingüística (término), comunicativa (contexto) y cognitiva (concepto) (Cabré, 2003). Por esta razón, son consideradas como unidades poliédricas que “pueden definirse así desde cada una de las tres caras, lo que explicaría la diversificación de acepciones del término según el tipo de especialista o especialidad que lo define” (Cabré, 1999, p.54), mas esas tres caras son inseparables del término.
- Los términos no forman parte de un sistema autónomo como lenguaje especializado. En otras palabras, la TCT plantea que los términos no están separados de las palabras. Ambos son unidades que forman parte del léxico de un hablante. De hecho, los términos están asociados a las unidades léxicas. Estas son unidades compuestas de un concepto y denominación y que cumplen funciones referenciales, expresivas o conativas. Por ende, en un principio estas unidades no son “ni palabras ni términos sino potencialmente términos o no términos” (Cabré, 1999, p.56).
- Por su parte, debido a su multidimensionalidad, “... [los términos] pueden ser analizados desde otras perspectivas y comparten con otros signos de sistemas no lingüísticos el espacio de la comunicación especializada” (Cabré, 1999, p.52). En efecto, esta función dentro de la comunicación especializada se activa según el contexto. Por ello, las unidades pasan a ser términos cuando están presentes ciertos rasgos morfosintácticos, semánticos y pragmáticos que le dan dicha clasificación dentro de un ámbito específico.
- Así como las unidades léxicas se componen de concepto y denominación, los términos poseen de igual manera una forma o denominación y un significado o contenido. A diferencia de lo que sostiene la TGT, el concepto no precede a la denominación, sino que ambas son simultáneas. De esta manera, el contenido y la forma están asociados, por lo que la unidad léxica no solo cumple una función denominativa, sino que contiene significado (Cabré, 2003).
- De igual manera, tanto en la TGT como en la TCT se plantea que el valor de un término viene determinado por la posición que ocupa en el sistema de conceptos. Sin embar-

go, en la TCT, a diferencia de en la TGT, los términos no pertenecen a un único ámbito especializado, pero “son usados en un ámbito con un valor singularmente específico” (Cabré, 1999, p.57).

- Se admite la variación de los conceptos y sus denominaciones. En otras palabras, no se sostiene una idea de control evolutivo. También, se considera que las unidades léxicas son polisémicas, pero se admite que algunas unidades pueden pertenecer solo a la comunicación general y aún no se han activado como términos, o que algunos términos solo se utilizan en un ámbito específico o que inclusive un término está presente en diversos ámbitos y en la comunicación general. Esto permite explicar la banalización (unidades especializadas que pasan a formar parte del lenguaje general), terminologización (unidades léxicas del lenguaje general que pasan al especializado) y la pluriterminologización (transferencia de unidades especializadas de un ámbito a otro). Se acepta la sinonimia como un fenómeno de la realidad comunicativa. Ahora bien, “cuanto más especializado es el texto mayor es su sistematicidad y menor su grado de variación denominativa” (Cabré, 1999, p.55).

Para Cabré (1999), el objetivo de la terminología desde el punto de vista teórico consiste en describir formal, semántica y funcionalmente las unidades terminológicas, también denominadas “términos”. Consideramos que este paradigma conviene particularmente al enfoque de nuestro trabajo, en que adoptamos una visión amplia del término, acorde con el contexto lingüístico y comunicativo en que tiene lugar la traducción especializada.

Así como la Teoría Comunicativa de la Terminología es un corpus de ideas original, que configura una nueva visión sobre la ter-

minología y sus unidades, desde un punto de vista metodológico, la TCT adopta principios y supuestos que proceden, mayoritariamente, de la lingüística descriptiva, y en particular, de la lingüística de corpus y de la lingüística del texto.

Por todo ello, la necesidad de adecuación a distintas necesidades y contextos socioprofesionales; el enfoque predominantemente descriptivo; el respeto a la variación terminológica (tanto denominativa como conceptual); y la necesidad de considerar la multidimensionalidad del término (dimensión cognitiva, lingüística y comunicativa) tiene una serie de consecuencias metodológicas sobre la identificación de las unidades, el valor terminológico, la delimitación de los términos, los ámbitos de especialidad, etc., que iremos viendo a lo largo de este trabajo.

2. Los escenarios terminológicos

Puesto que, de acuerdo con la TCT, las unidades terminológicas se pueden observar y utilizar en contextos muy diferentes, y con visiones diferentes, en este trabajo acuñamos el concepto de “escenario terminológico”. Un escenario terminológico se define por la intersección de tres dimensiones:

- a. el ámbito o contexto profesional donde se usa el término;
- b. el producto o aplicación que se quiere generar, y
- c. la finalidad que se quiere perseguir.

Consideramos que el planteamiento de cualquier trabajo terminológico debería partir por explicitar con claridad sus tres dimensiones, para poder definir posteriormente sus etapas metodológicas y herramientas de trabajo.

En la figura 1, ejemplificamos la idea de escenario terminológico.

La lista de ejemplos que proponemos no es exhaustiva, pero sí bastante amplia. En la columna de la izquierda, presentamos casos de ámbitos (o contextos) profesionales; en la columna del medio, posibles productos o aplicaciones en los que los términos ocupan un lugar relevante; en la columna de la derecha, ejemplos de la finalidad que se persigue con el producto generado, en cada contexto determinado. Los colores rojo, naranja y negro ejemplifican posibles escenarios concretos.

Ejemplos de escenarios terminológicos serían:

- Una compilación de un diccionario para traductores profesionales
- La creación de un diccionario normalizado de un ámbito determinado

c. Una recopilación de neologismos de prensa política venezolana

d. La creación de un glosario previo a una traducción especializada para traducir de forma colaborativa.

Este último es el escenario donde situamos nuestro trabajo.

Cada uno de esos escenarios conlleva un conjunto de prácticas, etapas metodológicas, procesos de trabajo y herramientas no necesariamente coincidentes.

3. El encargo de trabajo, las etapas metodológicas y las herramientas

Las reflexiones metodológicas que presentamos en este artículo se basan en un encargo de trabajo real, que la Comisión Económica para América Latina y el Caribe (en adelante, CEPAL) trasladó a los estudiantes del Magíster de Tra-

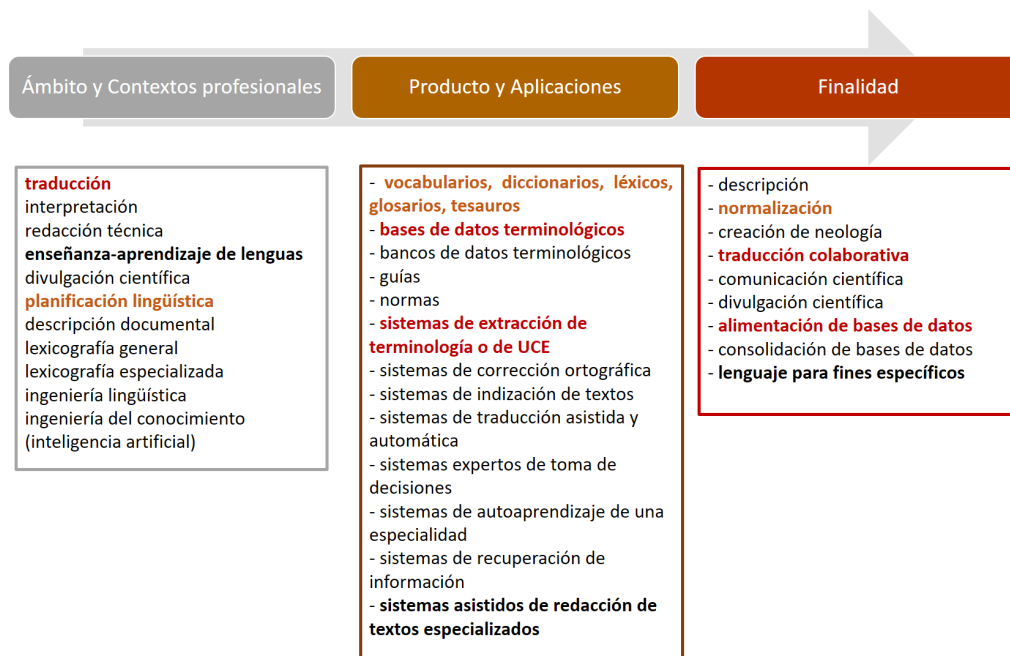


Figura 1. Escenarios terminológicos.

ducción Inglés-Español de la Pontificia Universidad Católica de Chile. Este encargo fue llevado a cabo en el curso de Gestión terminológica aplicada a la traducción, con la supervisión de sus profesores, que firman este trabajo como responsables, y de los expertos terminólogos, traductores y revisores de la CEPAL, en 2015.

El encargo decía así:

En nuestra región la pobreza persiste como un fenómeno estructural que caracteriza a la sociedad latinoamericana. Con el objeto de contribuir a un diseño más integral de las políticas públicas para superar la pobreza y la desigualdad socioeconómica, la CEPAL considera que hace falta analizar la pobreza desde una perspectiva multidimensional, y no solo desde el análisis tradicional de los ingresos medidos en PIB per cápita. El índice de pobreza multidimensional integra aspectos monetarios y no monetarios del bienestar.

Nos interesa seleccionar y definir términos económicos y relacionados con las estadísticas o índices de medición de la pobreza. Ojalá que los términos seleccionados sean nuevos, es decir, NO estén contenidos en la base de datos terminológica UNTERM.

El contexto del trabajo era la traducción especializada del área económica, y en concreto, de un ámbito temático emergente como la pobreza multidimensional; el producto, la creación de un glosario bilingüe inglés-español con términos de este ámbito; y la finalidad del glosario, establecer equivalentes terminológicos fiables entre ambas lenguas, con el fin de tener un recurso de referencia validado para las traducciones que la CEPAL lleva a cabo regularmente en este ámbito.

Con el escenario dibujado, procedimos a determinar las etapas metodológicas del trabajo:

1. Selección del corpus

2. Procesamiento con dos cadenas diferentes de herramientas de análisis y extracción terminológica
3. Exportación a una base de datos: glosario provisional
4. Contraste con UNTERM, generación de la nomenclatura definitiva
5. Análisis de la selección de unidades terminológicas
6. Introducción de sinónimos y variantes
7. Compleción y validación de las equivalencias inglés-español
8. Elaboración de definiciones en español
9. Revisión y compleción de las informaciones en inglés
10. Revisión final de la terminología CEPAL
11. Revisión final de la revisora CEPAL
12. Entrega del producto final

Las herramientas informáticas que utilizamos para este trabajo fueron, principalmente: alineadores, concordanciadores, sistemas de extracción de terminología, gestores de bases de datos terminológicos, y conversores de formatos.

4. Metodología: la etapa de procesamiento de los corpus

Las etapas metodológicas necesarias para cumplir con los requerimientos de este encargo —seleccionar y definir los nuevos términos económicos y relacionados con las estadísticas o índices de medición de la pobreza no contemplados en la base terminológica UNTERM—, se delimitan en dos grandes líneas: la primera, relacionada con el procesamiento y explotación del corpus paralelo, con la finalidad de extraer información relevante y proporcionar

una lista inicial de candidatos a término, y la segunda, de carácter más terminológico, de decisión humana sobre los contenidos ofrecidos por las herramientas de procesamiento.

A partir de los informes redactados en inglés y español sobre la pobreza multidimensional entregados por CEPAL, se procedió a su limpieza (eliminación de toda información no lingüística como tablas, gráficos, números de páginas,

etc.), con el fin de poder procesarlos con las herramientas de análisis lingüísticos. La primera etapa consistió en crear un bitexto, un texto paralelo elaborado a partir de un documento original y su traducción en que los segmentos de la lengua fuente y meta se hicieron corresponder paralelamente, como se puede observar en la figura 2.

Tabla 1. Planificación del trabajo terminológico.

Etapla 1 PROCESAMIENTO DE LOS CORPUS	Etapla 2 TRABAJO TERMINOLÓGICO
Selección y preprocesamiento del corpus paralelo (bilingüe)	Contraste con UNTERM, generación de la nomenclatura definitiva
Procesamiento con dos cadenas diferentes de herramientas de análisis y extracción terminológica	Análisis de la selección de unidades terminológicas
Exportación a una base de datos: glosario provisional	Introducción de sinónimos y variantes
	Compleción y validación de las equivalencias inglés-español
	Elaboración de definiciones en español
	Revisión y completación de las informaciones en inglés
	Revisiones finales con la CEPAL

STEPS IN T (...)	PASOS PARA (...)
Measuring (...)	DE LA POBR (...)
Accordingly, as with the monetary measure, the requisite information must be disaggregated at the individual or household level, such as that obtained from household surveys or population censuses.	Por tanto, de manera similar a la medición monetaria, se requiere información desagregada al nivel de las personas o los hogares, como la que se obtiene de las encuestas de hogares o los censos de población.
In practic (...)	En la prác (...)
Each of th (...)	Para todos (...)
Before selecting dimensions, indicators and thresholds, the normative approach underlying the measurement must be defined.	Para seleccionar las dimensiones, indicadores y umbrales, se requiere definir el enfoque normativo que subyace a la medición.
This is no easy task: first, because there are no conclusive arguments to claim that any given poverty approach is intrinsically superior to any others; and, second, because poverty is measured in a wide variety of settings.	Esta no es una tarea fácil, puesto que no hay argumentos suficientes para afirmar la superioridad intrínseca de algún enfoque de la pobreza por sobre los demás y por la amplia variabilidad entre los distintos contextos en los que se mide la pobreza.
This means (...)	Esto signi (...)
Accordingly, crucially important steps include choosing an approach that is relevant to the reality of poverty in a specific context, identifying potential practices for attaining the thresholds, and clarifying the feasibility of empirically testing whether people actually attain those thresholds.	Así, la selección de un enfoque pertinente a la realidad de la pobreza en el contexto, la determinación de las posibilidades prácticas de lograr los umbrales y el esclarecimiento de la viabilidad de comprobar empíricamente si las personas alcanzan dichos umbrales revisten importancia crítica.

Figura 2. Vista previa del bitexto elaborado con la herramienta YouAlign® a partir de los informes bilingües publicados por CEPAL.

Fuente: Disponible en <https://youalign.com>

Terminado este proceso, se creó un corpus paralelo, a partir del bitexto elaborado anteriormente, en Sketch Engine^{®1}, una plataforma para la gestión de corpus lingüísticos que integra un software de análisis textual que permite estudiar el comportamiento del lenguaje por medio de búsquedas efectivas (extractores de terminología, concordanciadores, etiquetadores, buscadores multilingües y paralelos, etc.) de acuerdo con el objetivo que se quiere alcanzar; en nuestro caso, la extracción de candidatos a términos y contextos definitorios relevantes.

Además de Sketch Engine[®], los estudiantes trabajaron con la plataforma Terminus 2.0^{®2}, una estación de trabajo específica para la terminología que incluye la cadena completa del trabajo terminográfico: constitución y exploración de corpus textuales, extracción de términos a partir de los corpus creados, creación de glosarios o bases de datos terminológicos, gestión y mantenimiento de glosarios y proyectos, y edición de glosarios son algunas de sus herramientas.

El trabajo con ambas herramientas se justifica debido a las especificidades de cada una, que serán descritas con mayor detalle en el apartado 7, Conclusiones. Para facilitar la visualización de los procesos llevados a cabo por los estudiantes, mostraremos ejemplos extraídos de las dos herramientas, según sus aplicaciones.

El primer paso metodológico fue, por lo tanto, extraer los candidatos a términos por medio de los extractores asociados a cada estación de trabajo y analizar tanto su frecuencia como su pertinencia para el área. El proceso consistió en contrastar el corpus de trabajo con un corpus de referencia integrado a la plataforma

para determinar, a partir de criterios estadísticos, las palabras simples o expresiones multipalabra que tienen mayor índice de terminológica dentro del corpus (figura 3).

Con Terminus 2.0[®], los candidatos a términos formados por expresiones multipalabra también se pueden extraer de acuerdo con patrones sintácticos previamente definidos, como se puede observar en la siguiente imagen (figura 4).

Esta primera extracción generó una lista de candidatos a términos que fueron, posteriormente, validados por los estudiantes, tarea que se pudo realizar a través de la concordancia de cada unidad de análisis de forma individual.

El concordanciador, asociado a la herramienta KIWC (*Key Word in Context*), que se refiere al texto resaltado en rojo en una concordancia, permitió analizar el contexto de ocurrencia del candidato a término y, a partir de este análisis, determinar no solo su pertinencia como término, sino buscar por contextos definitorios o explicativos correspondientes.

Durante el trabajo con el concordanciador se pudo observar que la herramienta incurrió en un error de silencio, que sucede cuando el programa no es capaz de delimitar el término adecuadamente. Un análisis manual pudo identificar el sustantivo “index” como parte del compuesto terminológico “multidimensional poverty index”. Este tipo de análisis se llevó a cabo para validar todos los candidatos a términos obtenidos por los sistemas de extracción de ambas herramientas. Esta lista inicial se exportó a una base de datos provisional en Excel, donde se registró toda la información pertinente.

Esta base de datos provisional estaba compuesta de 136 candidatos a términos, que fueron considerados el punto de partida para ge-

1 Disponible en <https://app.sketchengine.eu>

2 Disponible en <http://terminus.iula.upf.edu/cgi-bin/terminus2.0/terminus.pl>

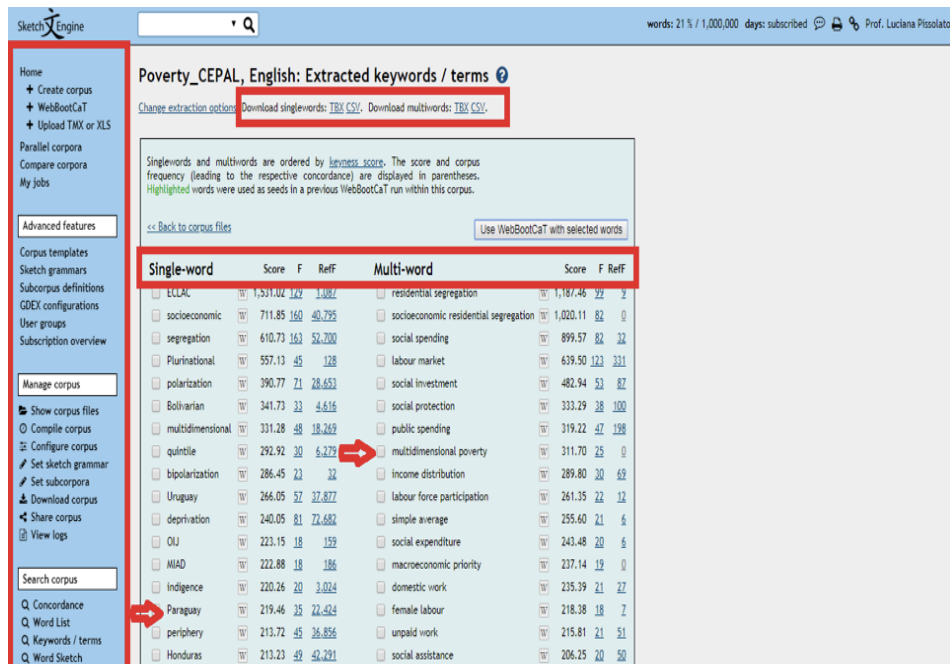


Figura 3. Extracción automática de terminología con Sketch Engine®.

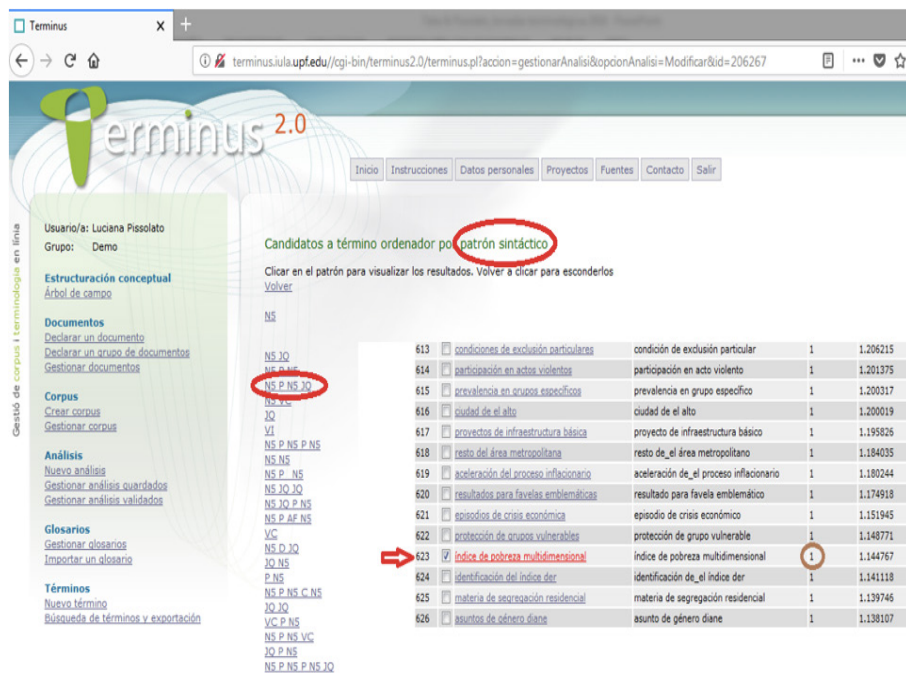


Figura 4. Extracción de candidatos a términos por patrón sintáctico [Nombre+Preposición+Nombre+Adjetivo], con Terminus 2.0®.

Reflexiones sobre la metodología terminológica en contextos de traducción: ¿qué se pierde en el trabajo con herramientas informáticas?



Figura 5. Concordancia del candidato a término "multidimensional poverty", con Sketch Engine®.

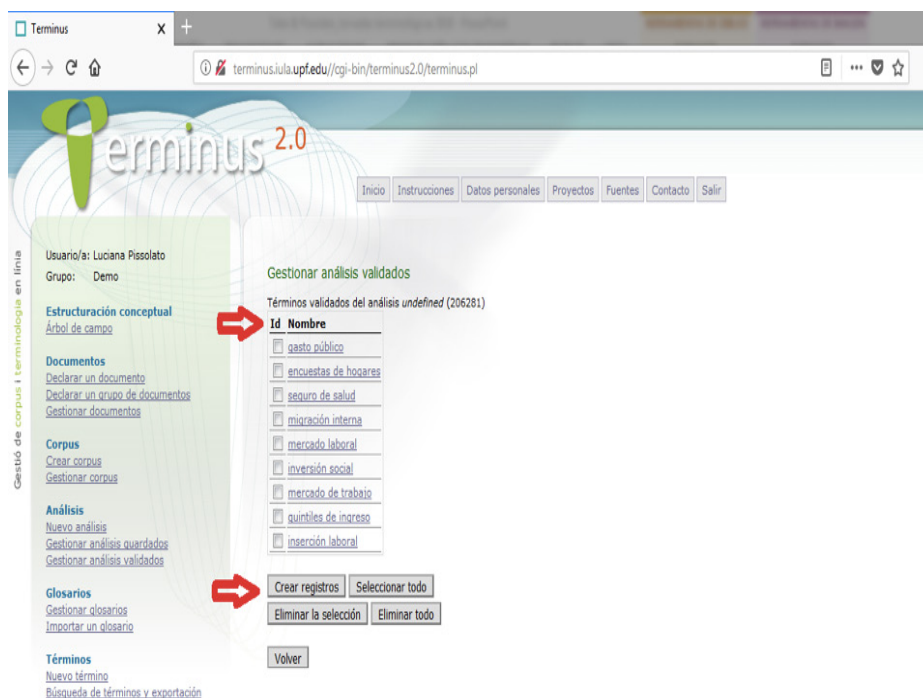


Figura 6. Creación de la base de datos final, elaborada en la plataforma Terminus 2.0®.

Terminus 2.0

Inicio | Instrucciones | Datos personales | Proyectos | Fuentes | Contacto | Salir

Cabecera

Usuario/a: Luciana Pissolato
Grupo: Demo

Glosario: Pobreza Cepal
Proyecto: Pobreza Cepal
Autor/a del término: Luciana Pissolato
Fecha de creación: 2018-08-28

Información básica

Id: 105206
Término: mercado laboral
Lengua: Castellano
Categoría gramatical: Nombre
Fuente del término: Cepal
Tipo de fuente: Texto especializado
Área temática: Economía
Estatus del término: Sin atribuir
Organismo normalizador: -
Lengua (Marca geográfica): Castellano

Definiciones

Fuente: -

Contextos

entre las mujeres que recurren al trabajo por cuenta propia se encuentran aquellas que han iniciado microempresas motivadas por la falta de oportunidades en el mercado laboral asalariado y por la necesidad de acceder a ingresos propios.

Fuente: Cepal

Equivalentes

Término	Lengua	Categoría gramatical	Fuente
no hay resultados.			

Remisiones

Término	Tipo de remisión	Fuente	Categoría gramatical
no hay resultados.			

Colocaciones

Colocación	Categoría gramatical	Posición del colocativo
no hay resultados.		

Modelo ficha terminológica Terminus 2.0

Figura 7. Model

nerar, en la siguiente etapa, la lista definitiva de términos.

En las dos figuras siguientes se puede observar la creación de la base de candidatos a términos, y el modelo de ficha terminológica creada en Terminus 2.0®.

Como pudimos observar, la metodología basada en corpus requiere de un proceso de triangulación de datos, fundamental para la extracción y posterior análisis de los datos recuperados por las herramientas y por el terminólogo, en su labor puntual.

La triangulación de datos, tal como la entendemos en este trabajo, se define como una “combination, in a integrated manner, of multiple (two or more) corpora in one study of a single phenomenon” (Malamatidou, 2018, p. 33), pero es también entendida en la literatura, de manera general, como la combinación de múltiples teorías, fuentes de datos, métodos o investigadores en el estudio de un único fenómeno.

El trabajo con diferentes plataformas de explotación y gestión de corpus permitió a los traductores, en su rol de terminólogos, hacer referencias cruzadas y corroborar la información recolectada sobre los términos, lo que les facilitó la validación de los datos obtenidos. Además, contar con instrumentos de triangulación de datos es de gran relevancia para esta investigación, puesto que combinar diferentes tipos de corpus, como el corpus comparable puesto a disposición por CEPAL y bancos de datos terminológicos internacionales, como UNTERM, constituyen fuentes complementarias de datos lingüísticos cuya integración a la labor terminológica promueve una calidad del trabajo más alta.

De esta manera, y por medio de un trabajo colaborativo entre los estudiantes, se pudo completar, al final de todo el proceso, el encargo proporcionado: seleccionar y definir los nuevos términos económicos, que fueron relacionados con las estadísticas o índices de medición de la pobreza, y que no estuvieron contemplados por UNTERM.

TÉRMINOS CANDIDATOS DEL
CORPUS SOBRE POBREZA

Término en inglés	Término en español	UNTERM
93. Multidimensional poverty measurement	Medición multidimensional de la pobreza	No
94. Mutidimensional poverty index	Índice multidimensional de pobreza	Si está el término en inglés, pero no se utiliza el equivalente en español en el texto paralelo.
95. Multidimensional poverty indicator	Indicadores multidimensionales de pobreza	No
96. Public policy	Políticas públicas	No
97. Rights-based approach	Enfoque de derechos	Si está el término en inglés, pero no se utiliza el equivalente en español en el texto paralelo.
98. Capabilities approach	Perspectivas de capacidades	No
99. Human dignity	Dignidad humana	Sí
100. Citizenship	Ciudadanía	No
101. Welfare	Bienestar	Sí

Figura 8. Muestra de la base de datos provisional, contrastada con UNTERM.

5. Metodología: la etapa terminológica

A partir de la lista de 136 candidatos a términos con que concluía la etapa de procesamiento, se dio inicio a la segunda etapa del trabajo, dedicada al trabajo terminológico propiamente dicho, de decisión humana sobre los contenidos ofrecidos por las herramientas de la primera etapa.

Este segundo bloque de trabajo se dividió en distintas subetapas metodológicas, que ya fueron enumeradas en el apartado 3 (mantenemos la numeración de aquel apartado):

4. *Contraste con UNTERM, generación de la nomenclatura definitiva*
5. *Análisis de la selección de unidades terminológicas*
6. Introducción de sinónimos y variantes
7. Compleción y validación de las equivalencias inglés-español
8. Elaboración de definiciones en español

9. Revisión y compleción de las informaciones en inglés

10. *Revisión final de la terminología CEPAL*

11. *Revisión final de la revisora CEPAL*

12. Entrega del producto final

Puesto que, por razones de espacio, no es posible presentar todas las decisiones tomadas en cada una de las subetapas, hemos marcado en cursiva aquellas sobre las cuales vamos a detallar algunos problemas metodológicos que encontramos en ellas y, sobre todo, presentar las primeras reflexiones en torno al contraste entre los datos arrojados por las herramientas de procesamiento la etapa anterior y el trabajo terminológico de razonamiento humano.

5.1. Contraste con UNTERM, generación de la nomenclatura definitiva

Esta etapa consistía en cumplir con uno de los requisitos exigidos por el encargo de la CEPAL,

que pretendía aplicar un primer filtro a la lista de partida: excluir de la nomenclatura inicial aquellas unidades terminológicas que ya se encontraban recogidas en la base de datos terminológicos UNTERM. Es decir, recoger solamente términos (y por lo tanto conceptos) nuevos, no descritos ni tratados anteriormente en la base de datos que la CEPAL utiliza como material de referencia obligado para sus traductores y terminólogos.

Esta etapa, que a priori debía resolverse término por término con una solución binaria (está/no está, sí/no), arrojó resultados inesperados, que hemos agrupado en tres tipos de problemas diferentes.³

5.1.1 Disparidad de usos terminológicos entre la traducción y UNTERM

Se observó, al contrastar la lista de términos extraídas del corpus con la base de datos de referencia (UNTERM) que, en muchas ocasiones, no se utilizaron los equivalentes normalizados por la institución en los documentos redactados y traducidos por CEPAL, que constituían nuestro corpus. Es el caso del término “income poverty”, que se encuentra registrado en las bases de UNTERM como “pobreza económica”, como se pudo visualizar en las siguientes imágenes. La figura 9 muestra la concordancia del término en Sketch Engine®, procedente del corpus, y la figura 10 las equivalencias declaradas en la base de datos UNTERM.

La misma incoherencia se observó en otros muchos casos. Por ejemplo, “gender inequality” aparecía traducido en el corpus (y recogido en nuestra lista provisional) como “desigualdad de género”, mientras que UNTERM

priorizaba “desigualdad por razón de sexo”. Otro caso: “global financial crisis” aparecía traducido en la lista como “crisis financiera global”, mientras que UNTERM establece que el equivalente consensuado en Naciones Unidas es “crisis financiera mundial”.

La primera observación inesperada entre las equivalencias que proporcionaba el glosario extraído del corpus de la CEPAL y la consulta manual en UNTERM fue que, en un número no menor de casos, los redactores y traductores de la CEPAL no utilizaban los términos recomendados por Naciones Unidas. Y detrás de esos casos puntuales, se observó el desconocimiento de algunas decisiones terminológicas y traductorales que podían proyectarse sobre muchas otras unidades terminológicas complejas o sintagmáticas que contenían esas unidades simples objeto de debate: *gender* > género/sexo, *global* > global/mundial.

5.1.2 Errores de sentido

Otra de las consecuencias inesperadas de este chequeo fue la detección de errores de sentido en la traducción de algunos de los textos del corpus. Por ejemplo, para “care work” aparecía en nuestro listado “trabajo de cuidado”, que obviamente es una equivalencia errónea: UNTERM fijaba como equivalentes “prestación de cuidados”, “trabajo doméstico” o “trabajo asistencial” según los contextos. “Multidimensional poverty index” aparecía traducido como “índice multidimensional de pobreza”, cuando es “índice de pobreza multidimensional”.

5.1.3 Diferencias entre la forma canónica del término y su uso discursivo

Otro grupo interesante de discordancias entre la lista de términos extraída del corpus y las consultas a UNTERM fue constatar las diferencias existentes a veces entre un término ais-

3. Las consultas a UNTERM fueron realizadas en mayo de 2016. La base de datos terminológica de Naciones Unidas se actualiza regularmente.

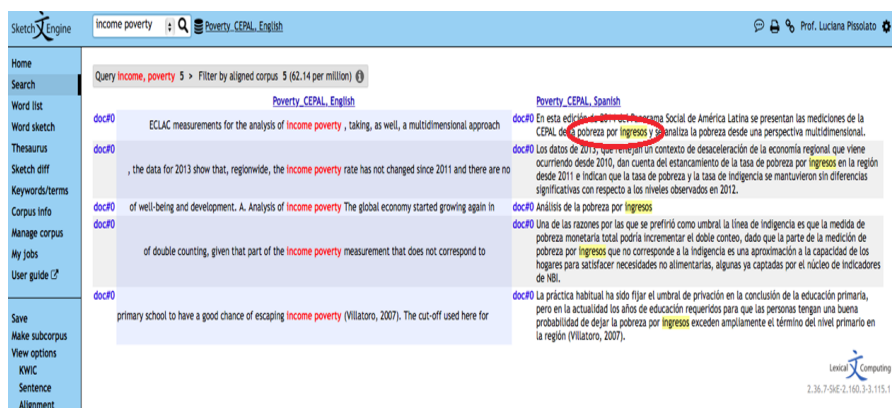


Figura 9. Concordancia de “income poverty” con Sketch Engine®.

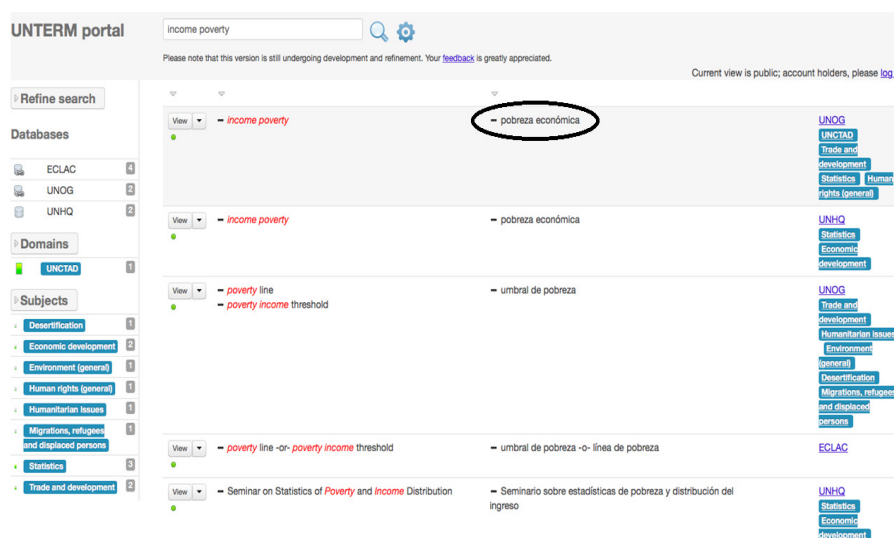


Figura 10. Consulta a UNTERM sobre el término “income poverty”.

lado en una base de datos, que se registra en su forma más lexicalizada, y su uso discursivo, que es el que se registra en su uso real, y que teníamos recogido en el listado. Esa constatación nos permitió constatar diferencias entre “structural poverty”, que había sido traducido por “pobreza como fenómeno estructural”, y “pobreza estructural”, el término de UNTERM. O “poverty reduction strategy”, recogido como “estrategia tendiente a superar la pobreza” versus “estrategia de reducción de pobreza”. En estos casos, consideramos que esta variación

terminológica de tipo discursivo es relevante para un traductor, y que debería recogerse en UNTERM, junto a las formas canónicas.

5.2 Análisis de la selección de unidades terminológicas

En la etapa 5, relativa al análisis de la selección de unidades terminológicas, los problemas se relacionaron con diferentes aspectos:

5.2.1 Delimitación terminológica

Tanto Sketch Engine® como Terminus 2.0® incurrieron en el error de silencio, que se define como la no selección de unidades terminológicas relevantes o —de manera más específica y aplicada al entorno tecnológico utilizado para este trabajo—, como la no delimitación correcta del término, caso de “multidimensional poverty” > “multidimensional poverty index”, como se presentó anteriormente en la figura 2 y como se puede observar también en la siguiente imagen, sobre el término “umbral de pobreza”, delimitado aquí como unidad terminológica simple pero, a partir de la extracción de su concordancia, se observa su carácter sintagmático:

Este tipo de error es común a todas las herramientas de extracción terminológica. Conforme pondera Estopà (2001, pp. 227-228), los sistemas de extracción de unidades terminológicas suelen generar demasiado silencio y ruido; proponer demasiados candidatos a términos que el usuario debe validar manualmente; se centran exclusivamente en la unidad terminológica entendida como una unidad

nominal y, la mayoría de ellos, se basan en la extracción de un tipo de unidad terminológica —la unidad terminológica poliléxica—; utilizan patrones para detectar estas unidades y se centran en las mismas técnicas para detectar y extraer todas las unidades.

5.2.2 Pertinencia del término para el área

Muchos candidatos seleccionados por las herramientas no fueron validados como términos en un posterior análisis manual puesto que no reflejaban un concepto económico relacionado con la medición de la pobreza, sino que se asociaban a otras áreas de especialidad o eran, más bien, propios de la economía general o de ámbitos próximos (sociología, políticas públicas, etc.); por esta razón, no fueron considerados para el glosario final. Es el caso de “coeficiente de endeudamiento”, por ejemplo. En otros casos, como “clase media”, hubo que recurrir posteriormente a los expertos de la CEPAL para discutir si efectivamente tenía un sentido específico en el ámbito de la pobreza multidimensional que justificara su selección en el glosario final.



Rank	Forma	Lema	Frecuencia	Ponderación
251	umbral	umbral	32	174.194360
252	áreas metropolitanas	área metropolitana	8	173.809205
253	gasto social	gasto social	104	173.511531
254	aproximación	aproximación	24	170.480511
255	what	what	4	169.561160
256	ipas	ipas	4	169.341452
257	discriminación	discriminación	37	169.339330
258	inei	inei	4	168.068564
259	equidad de género	equidad de género	2	167.543789
260	pdf	pdf	15	166.534632
261	ener	ener	8	166.509677
262	pánico urbano	pánico urbano	1	165.430894
263	enfoque de derechos	enfoque de derecho	3	164.706121
264	básicas	básico	16	164.671158

Figura 11. El problema de la delimitación terminológica. Extracción hecha con Terminus 2.0®.

Este tipo de problema es bastante frecuente cuando se trabaja con herramientas de extracción terminológica, y se considera un error de ruido. Tal error se debe, especialmente, a la imposibilidad de trabajar con un corpus de referencia elaborado por el propio investigador, ya que ambas herramientas lo traen incorporado. El hecho de no poder controlar el corpus de referencia, con el cual se contrasta el corpus de trabajo para la extracción de terminología, interfiere directamente en la calidad del resultado, puesto que los corpus de referencia incorporados a las plataformas son, mayormente, de lengua general (Sketch Engine®) o de áreas de especialidad previamente definidas (Terminus 2.0®) y esa peculiaridad imposibilita una selección inequívoca de las unidades de mayor índice de terminologicidad.

5.2.3 Selección de plurales lexicalizados

Este tipo de error ocurre cuando hay una disparidad entre la forma canónica y el uso discursivo que se le da a los términos. Se relaciona, frecuentemente, con problemas de lematización: la conversión de las formas flexionadas de las unidades léxicas en sus formas canónicas, o lemas. Tanto Sketch Engine® como Terminus 2.0®

disponen de un lematizador incorporado a su plataforma, lo que minimiza la selección de plurales lexicalizados; sin embargo, extraen ocurrencias de sintagmas pluralizados, lo que promueve la selección de tales unidades debido a su alta productividad dentro del corpus de trabajo. Es el caso de “índices de Atkinson”, “índices de polarización”, “coeficientes de exposición”, “encuestas de hogares”, etc. Este fenómeno pone en evidencia la diferencia de su forma canónica con el uso discursivo que se les da a las unidades terminológicas.

5.3 10/11 Revisión final de la terminología y la revisora de la CEPAL

Más brevemente, en la etapa final de revisión por parte de la CEPAL, tanto por parte de la experta terminóloga como de la revisora, surgieron algunas dudas y se hicieron algunas correcciones al glosario final que no se correspondían con los datos que se habían extraído de las herramientas en la primera etapa, y que se habían validado posteriormente en el trabajo terminológico manual.

Un grupo de esas correcciones tuvieron que ver con temas de coherencia entre el término inglés



Rank	Forma	Lema	Frecuencia	Ponderación
404	☐ índice de recuento	índice de recuento	17	113.414816
405	☐ índices de polarización	índice de polarización	11	113.298447
406	☐ estigma	estigma	19	113.277395
407	☐ measure	measure	3	113.193663
408	☐ abuso de alcohol	abuso de alcohol	1	112.398921
409	☐ urbanización	urbanización	13	112.282590
410	☐ cohesión	cohesión	23	112.206656
411	☐ desafío	desafío	19	112.156964
412	☐ regresión	regresión	9	111.818909
413	☐ cuaderno	cuaderno	1	111.701983
414	☐ edades	edad	19	111.430939
415	☐ índice de wolfson	índice de wolfson	11	111.247313
13	☐ gráfico	gráfico	275	2699.903686
14	☐ recursos	recurso	124	2564.547002

Figura 12. El problema de los plurales lexicalizados. Extracción hecha con Terminus 2.0®.

(documentado efectivamente con una forma *x*) y el término español (documentado con una forma *y* que no era tan homónima como debía esperarse). Al tratarse de un vocabulario en gran parte nuevo y en proceso de estabilización, se decidió de forma conjunta realizar algunas modificaciones a algunos términos para ganar en coherencia formal y semántica. Ejemplos de esas disonancias formales son “precarious social protection arrangements” vs “protección social precaria”, “school dropout among adolescents” vs “deserción adolescente”, “socioeconomic variable” vs “variable de segmentación socioeconómica”, “adjusted headcount ratio” vs “índice de recuento ajustado de pobreza”, etc.

Otro grupo de correcciones consistió en añadir variantes terminológicas por analogía con términos similares... aunque esas variantes, *stricto sensu*, no estaban documentadas en el corpus. Por ejemplo, “gender income gap” tenía como equivalente documentado “brecha salarial entre hombres y mujeres”, y se decidió añadir “brecha salarial por razón de sexo”. “Hazardous work” era “trabajo riesgoso”, y se añadió “trabajo de riesgo” y “trabajo peligroso”.

En la siguiente sección, describiremos algunas de las principales ventajas y desventajas observadas durante la etapa de extracción terminológica y delinearemos algunas de sus consecuencias para el trabajo del traductor, de manera global.

6. ¿Qué se gana y qué se pierde con el uso de herramientas versus la extracción manual?

6.1. ¿Qué se gana con el uso de herramientas?

6.1.1 Tratamiento de grandes volúmenes de datos

Una de las ventajas más relevantes del uso de tecnologías para el tratamiento de datos lingüís-

tico es el hecho de que, sin ellas, no sería posible explorar y analizar corpus textuales tan extensos y, en consecuencia, la calidad de los resultados y las posibilidades de abstracciones y generalizaciones que se podría hacer sobre el lenguaje no sería representativo ni tampoco fiable. Además, los corpus, tanto comparables como paralelos, permiten a los traductores hacer observaciones sistemáticas sobre el lenguaje, lo que tiene consecuencias directas para el desarrollo de competencias fundamentales para el profesional de la traducción, lo que discutiremos más detalladamente en el apartado 7, Conclusiones.

6.1.2 Diferentes posibilidades de análisis lingüísticos y aplicación de criterios sistemáticos

La incorporación de herramientas y métodos de la lingüística de corpus en los estudios de traducción especializada abre un abanico de posibilidades de extracción de información relevante tanto para la elaboración de recursos de apoyo —glosarios, bases de datos terminológicos, etc.— y para la labor de traducción como tal, como también para el análisis de traducciones previas con el fin de preparar una traducción. Por medio del uso de los concordanciadores, extractores de terminología mono y multilingüe, fraseología especializada, etc., los traductores pueden acceder a información específica y aplicarla directamente a los propósitos de su trabajo.

Otros aportes de las herramientas constituyen: la función de búsqueda —por frecuencia, por orden alfabético, por patrón sintáctico, entre otros—, la anotación gramatical automática y la aplicación de filtros —búsqueda por clases gramaticales específicas como gerundios, verbos, etc.—, la triangulación con otros bancos de datos —Google Translator®, Wikipedia®— y, finalmente, la posibilidad de reutilización por parte de otros investigadores, interesados en temas relacionados.

6.1.3 Formatos de entrada y de salida compatibles con herramientas de traducción asistida por computador (TAC)

Sketch Engine®, por ejemplo, acepta diferentes formatos de archivos de entrada para la creación de 1) corpus paralelos (.xls, y .tmx), 2) corpus monolingües (.doc, .docx, .htm, .html, .ods, .pdf, .tar.bz2, .tar.gz, .tei, .tgz, .tmx, .txt, .vert, .xlf, .xliff, .xls, .xlsx, .xml, .zip.), y para 3) la descarga de resultados de búsquedas efectivas al interior de la plataforma (.cvs, .tbx). Esta característica hace que el trabajo previo de compilación y limpieza de textos que conformarán el corpus de trabajo sea mucho más expedito.

Por otra parte, la posibilidad de descargar los resultados de las exploraciones realizadas y subirlas directamente a las herramientas TAC facilita sobremanera el trabajo de documentación y de organización de recursos. Es el caso, a modo de ilustración, de una lista de términos en formato .tbx, que puede ser convertida y fácilmente incorporada al gestor de terminología Multiterm®, asociado a la plataforma de traducción de SDL Trados Studio®.

6.2. ¿Qué se pierde con el uso de herramientas?

6.2.1 Problemas de extracción de información y fiabilidad de los resultados

Como mencionamos anteriormente, las herramientas de extracción terminológicas incurren frecuentemente en errores de ruido y de silencio, lo que se convierte en un problema —no menor— para el traductor, puesto que la necesidad de revisión manual es imprescindible. Por lo tanto, aunque las ventajas de su uso sean evidentes, el proceso de extracción debe ser considerado siempre como semiautomático.

6.2.2 Actualización de los corpus de referencia y fluctuación de los resultados

Muchas herramientas, tales como Sketch Engine®, trabajan con corpus de referencia previamente incorporados, sobre los cuales se permite poca o ninguna gestión. Además, estos corpus sufren constantes actualizaciones, puesto que son, casi siempre, corpus abiertos. Esto hace que los resultados obtenidos de cada extracción, en el tiempo, pueda variar.

6.2.3 Explotación de corpus paralelos

Otra desventaja observada es que no todas las herramientas disponibles para los traductores cuentan con recursos para la explotación de corpus paralelos, como es el caso de Terminus 2.0®. En estos casos, es necesario aplicar todas las etapas metodológicas con ambos corpus por separado, lo que implica una mayor inversión de tiempo por parte del traductor y que, en ocasiones, inviabiliza su uso.

6.3. ¿Qué se gana con la extracción manual?

6.3.1 Detección de usos y sentidos terminológicos

Como hemos visto en apartados anteriores, las herramientas aplicadas a la terminología trabajan con criterios de extracción de base fundamentalmente estadística. Sin embargo, debido al carácter multidimensional de los términos, el criterio estadístico no es suficiente para identificar aspectos más sensibles tales como la metáfora y la variación, o para detectar la correcta delimitación de un sintagma terminológico formado por varias unidades léxicas.

Si consideramos, conforme mencionamos anteriormente, que las unidades solo pasan a tener estatus de términos cuando se cumplen determinados criterios morfosintácticos, semánticos y pragmáticos, los cuales no están contemplados (por lo menos hasta la fecha)

por las herramientas de explotación de corpus, queda evidente la necesidad de la aplicación del criterio del investigador, quien da la palabra final sobre el estatus y delimitación de una unidad terminológica.

6.3.2 Adecuación y pertinencia al escenario

Aunque las herramientas puedan detectar correctamente las unidades terminológicas relevantes a partir de un corpus, identificar la adecuación a un área o subárea de especialidad y la pertinencia al escenario —en nuestro caso, seleccionar “términos económicos y relacionados con las estadísticas o índices de medición de la pobreza”, que “ojalá sean nuevos” y “NO estén contenidos en la base de datos terminológica UNTERMPortal”—, requiere de un trabajo de triangulación de datos, tal como lo que presentamos en la Metodología, que ninguna herramienta es capaz (todavía) de ofrecer.

6.3.3 No hay ruido

Finalmente, el trabajo de extracción terminológica manual, que se hace por medio de la lectura del material y de la identificación de la terminología basada en criterios morfosintácticos, semánticos y pragmáticos (Cabré, 1999), garantiza una adecuada recopilación de términos, pertinentes tanto al escenario (la creación de un glosario previo a una traducción especializada para traducir de forma colaborativa) como al área de especialidad (términos económicos relacionados con las estadísticas o índices de medición de la pobreza). En este sentido, no habrá ruido, o sea, el traductor, especialmente si cuenta con formación en terminología, (muy probablemente) no incidirá en el error de seleccionar unidades poco relevantes o, incluso, irrelevantes para el encargo. Quizás, producto de la naturaleza manual del

trabajo, incurra en el error de silencio: no selección de términos relevantes.

6.4. ¿Qué se pierde con la extracción manual?

Por último, consideramos que tres son las desventajas más importantes del trabajo de extracción manual: el tiempo, del cual no siempre dispone el traductor; la subjetividad en relación con los criterios para la selección de las unidades terminológicas, que pueden variar entre uno y otro profesional; y la no exhaustividad, aunque no siempre es necesario que un trabajo recopile integralmente todos los términos presentes en un corpus.

7. Conclusiones: implicaciones de la metodología basada en corpus para la formación de traductores

Las herramientas de explotación y gestión de corpus, tales como Sketch Engine® y Terminus 2.0®, son recursos que colaboran de manera efectiva en la adquisición de la competencia en traducción, tanto en lo que respecta a la resolución de problemas puntuales —como es el caso de la terminología, sobretodo en áreas interdisciplinarias—, como en la adquisición de competencias para el desarrollo de estrategias eficientes para la utilización de los recursos y herramientas disponibles para la traducción especializada.

A partir de los años 90, en la literatura, se ha discutido mucho sobre el papel de los diferentes tipos de corpus en la enseñanza de la traducción. Según Zanettin (1998), el uso de corpus paralelos ayuda a fortalecer la comprensión de los estudiantes de los textos fuente y a mejorar su expresión en la lengua meta; Wang (2004a) señala que, además, permite a los estudiantes desarrollar conciencia, reflexión y reacción,

características que distinguen a los traductores profesionales de los amateurs. Además, el uso de corpus revela las estrategias y métodos de traducción adoptados por traductores profesionales (Pearson, 2003). Bernardini *et al* (2007), subrayan aún que los corpus paralelos son útiles para mejorar la autonomía y la flexibilidad en traducción, además de brindar oportunidades de desarrollar en los estudiantes habilidades como innovación y resolución de problemas.

Todo ello contribuye sustancialmente para el desarrollo de competencias indispensables para el quehacer del traductor profesional.

Sin embargo, muchos estudios que asocian la lingüística de corpus a la traducción especializada se enfocan, especialmente, en la evaluación del producto. De acuerdo con Rodríguez-Inés (2017):

Although many corpus-based studies have focused on just the product of translation, i.e. translated texts, especially in terms of their quality, there are not so many corpus-based studies in which both the process and the product have being taken into account (p. 244).

Lo que intentamos hacer con la metodología aplicada en este trabajo es facilitar el desarrollo de estrategias en los estudiantes que les permita mejorar su proceso de traducción y, por ende, mejorar la calidad de su trabajo.

Uno de los objetivos de aprendizaje dentro de un currículum en traducción es comprender cómo utilizar los gestores de corpus y sus diferentes aplicaciones para la traducción especializada. En una primera etapa, como la que describimos en este trabajo, el principal objetivo fue el estudio de los términos en su entorno natural —corpus paralelo elaborado a partir

de informes sobre la pobreza de CEPAL— para la elaboración de un glosario basado en una necesidad real de la institución y que trae implicancias directas para la homogeneidad de las futuras traducciones de la institución. Para lograr este objetivo, el uso de herramientas de análisis de corpus lingüísticos es fundamental, ya que, conforme pondera Hu (2018, p. 183) dos de sus mayores ventajas son 1) contar con la extracción y análisis automáticos de los datos y 2) disponer de abundantes ejemplos de traducciones, realidades que validan el trabajo de búsqueda de equivalencias terminológicas.

Además de la aplicación mencionada anteriormente, según Corpas-Pastor (2012):

La utilización de corpus paralelos alineados permite a los traductores en formación estudiar las estrategias empleadas por los traductores profesionales para resolver tal o cual problema, comprobar las pérdidas y alteraciones de la información que se producen en el contexto de la comunicación mediada, y, a partir de ahí, empezar a desarrollar sus propias estrategias futuras. Sin embargo, y a pesar de su potencial, es un hecho que los corpus paralelos no se emplean con fines pedagógicos tan frecuentemente como los comparables (p. 11).

Una segunda etapa —que suele asociarse con los trabajos de conclusión de curso, de carácter investigativo—, comprende una descripción más amplia del lenguaje, impulsada por las herramientas de la lingüística de corpus, tales como: identificación de patrones de uso, fraseologías, variaciones terminológicas, frecuencia de palabras conceptualmente relevantes, características cohesivas de los textos especializados, macro y microestructuras, entre otros rasgos (Vila Barbosa, 2013), y tales trabajos contribuyen tanto a los estudios de los procesos como de los productos de la traducción especializada, y apuntan a formular nuevas propuestas dentro del área.

Referencias

- Austermühl, F. (2001). *Electronic Tools for Translators*. Londres: Routledge.
- Bernardini, S., Steward, D. y Zanettin, F. (2007). Corpora in translator education: an introduction. En F. Zanettin, S. Bernardini y D. Steward (eds.), *Corpora in Translation Education* (pp.1-14). Beijing: Foreign Language Teaching and Researching Press.
- Cabré, M. T. (1999). *La terminología. Representación y comunicación*. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.
- Cabré, M. T. (2003). Theories of Terminology. Their Description, Prescription and Explanation. *Terminology*, 9(2), 163-199. DOI: <https://doi.org/10.1075/term.9.2.03cab>
- Cabré, M. T. y Estopà, R. (2005). *Unidades de conocimiento especializado: caracterización y tipología*. En M. T. Cabré y C. Bach (eds.) *Coneixement, llenguatge i discurs especialitzat*, Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.
- Corpas Pastor, G. (2012). Corpus, tecnología y traducción. En M. Casas Gómez (dir.) y M.^a García Antuña (ed.), *XII Jornadas de Lingüística*. Cádiz: Servicio de Publicaciones de la Universidad.
- Estopà, R. (1999). Eficiencia en la extracción automática de terminología. *Perspectives*, 7(2), 277-286. DOI: <https://doi.org/10.1080/0907676X.1999.9961365>
- Estopà, R. (2001). Extracción de terminología: elementos para la construcción de un extractor. *Tradterm*, 7, 225-250. DOI: <https://doi.org/10.11606/issn.2317-9511.tradterm>.
- Freixa, J. (2005). Variación terminológica: ¿Por qué y para qué? *Meta: Journal des traducteurs*, 50(4). DOI: <https://doi.org/10.7202/019917ar>
- Hu, K. (2018). *Introducing Corpus-based Translation Studies*. New York: Springer.
- Malamatidou, S. (2018). *Corpus Triangulation. Combining Data and Methods in Corpus-Based Translation Studies*. Londres: Routledge.
- Oster, U. (2005). El traductor técnico y los términos: una vieja relación vista desde nuevas perspectivas. En: M. L. Romana García (coord.), *Actas del II Congreso Internacional de la Asociación Ibérica de Estudios de Traducción e Interpretación*, Madrid.
- Pearson, J. (2003). Using Parallel Texts in the Translator Training Environment. En F. Zanettin, S. Bernardini y D. Steward (eds.), *Corpora in Translator Education* (pp. 15-24). Manchester: St Jerome.
- Rodríguez-Inés, P. (2017). Analysis of the Translation Competence corpus from PACTE's experiment. En A. Hurtado Albir (ed.), *Researching Translation Competence by PACTE Group* (pp. 243-268). Amsterdam: John Benjamins.
- Tebé, C. (2006) *La representació conceptual en terminologia: l'atribució temàtica als bancs de dades terminològiques*. Editorial: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra (Sèrie Tesis, 15). Barcelona: IULA.
- Vargas Sierra, C (2002). *Utilización de los programas de concordancias en la traducción especializada*. Comisión Europea, Servicio de Traducción.
- Vila Barbosa, M. (2013). Corpus especializados como recurso para la traducción: análisis de los marcadores de la cadena temática en artículos científicos sobre enfermedades neuromusculares en pediatría. *Onomázein*, (27), 78-100.
- Wang, K. (2004a). *A parallel corpus: research and application*. Beijing: Foreign Language Teaching and Researching Press.

Wang, K. (2004b). The use of parallel corpora in translator training. *Computer-assisted Foreign Language Education*, 6, 27-32.

Zanettin, F. (1998). Bilingual comparable corpora and the training of translators. *Meta: Journal des traducteurs*, 43(4), 616-630. DOI: <https://doi.org/10.7202/004638ar>

Cómo citar este artículo: Tebé, C. y Pissolato, L. (2019). Reflexiones sobre la metodología terminológica en contextos de traducción: ¿qué se pierde en el trabajo con herramientas informáticas? *Mutatis Mutandis*, 12(1), 272-293. DOI: 10.27533/udea.mut.v12n1a11