

Nuevas perspectivas de análisis e investigación en fraseología: la aportación de los corpus y de la dimensión computacional

Giuseppe Trovato
Università Ca' Foscari Venezia, Italia

Abstract This note critically reviews a recently published volume, *Sistemas fraseológicos en contraste. Enfoques computacionales y de corpus* (2021) by Gloria Corpas Pastor, María Rosario Bautista Zambrana and Carlos Manuel Hidalgo-Ternero from the University of Malaga. It is a collection of essays written by specialists in the above-mentioned field, which highlights the increasing importance of the relation between Phraseology, Corpus Linguistics and Computational Linguistics. In fact, the interweaving of Phraseology with Corpus Linguistics, Natural Language Processing and Computational Linguistics has undoubtedly meant a real revulsive for the discipline.

Keywords Phraseology. Corpus Linguistics. Computational approaches. Linguistic technologies. Re-search.



Edizioni
Ca' Foscari

Submitted 2022-02-07
Published 2022-06-22

Open access

© 2022 | Creative Commons Attribution 4.0 International Public License



Citation Trovato, G. (2022). "Nuevas perspectivas de análisis e investigación en fraseología: la aportación de los corpus y de la dimensión computacional". *Rassegna iberistica*, 45(117), 147-154.

DOI 10.30687/Ri/2037-6588/2022/01/008

1

Comares es una prestigiosa editorial de Granada y especializada, entre otras disciplinas, en cuestiones lingüísticas y traductológicas en el marco de la colección «Interlingua». Tiene, efectivamente, una tradición consolidada en la publicación de volúmenes monográficos y colectivos sobre numerosos aspectos de la lingüística aplicada, contrastiva, la didáctica de las lenguas extranjeras, la traducción y la interpretación. El volumen *Sistemas fraseológicos en contraste: enfoques computacionales y de corpus* se enmarca de lleno en los ámbitos antes mencionados, poniendo especial atención en las nuevas perspectivas de análisis e investigación del fenómeno fraseológico (cf. Corpas Pastor 1996; 2000; 2003), concebido como el conjunto de modismos, frases hechas, locuciones figuradas, metáforas y comparaciones fijadas, refranes existentes en una determinada lengua, en el uso individual o en el de algún grupo. En efecto, es posible afirmar, sin temor a equivocarse, que la fraseología se configura como una parcela de la lingüística aplicada encaminada a estudiar una serie de unidades de sintaxis total o parcialmente fijas. El volumen aborda, pues, la importancia de los enfoques computacionales y de corpus (Corpas Pastor, Mitkov 2019; Corpas Pastor, Colson 2020; Mitkov et al. 2018) aplicados al estudio de los sistemas fraseológicos en contraste y es editado por especialistas en la materia, a saber, Gloria Corpas Pastor, María Rosario Bautista Zambrana y Carlos Manuel Hidalgo-Ternero, todos de la Universidad de Málaga.

Cuando la fraseología concebida como disciplina con una relevancia científica autónoma empezó a afirmarse en el panorama de la investigación, nadie hubiese imaginado los pasos que habría dado en menos de tres décadas y a partir de estudios pioneros como los Casares (1992), Corpas Pastor (1996), Martínez Marín (1996), Ruiz Gurillo (1998), Tristán Pérez (1988) y Zuluaga Ospina (1980), por mencionar los más paradigmáticos.

2

En la introducción, los autores señalan el nuevo rumbo que los estudios fraseológicos han venido tomando en los últimos tiempos y apuntan a la imbricación de la fraseología con la lingüística de corpus, el procesamiento del lenguaje natural y la lingüística computacional. Este nuevo fenómeno en el marco de la investigación fraseológica implica claramente nuevas perspectivas de análisis y se caracteriza por cierta hibridación metodológica.

Los capítulos se vertebran en torno a tres grandes bloques. El primero hace hincapié en aspectos considerados problemáticos para la fraseología contrastiva y, particularmente, para la computacional.

Abre el volumen el estudio de Silvia Cataldo titulado «La desautomatización fraseológica como fuente de humor: propuesta de clasificación e implicaciones para la traducción entre italiano y español». La autora combina el fenómeno lingüístico-cultural del humor con el de las unidades fraseológicas con valor metafórico, apuntando al fenómeno de la manipulación. Sus reflexiones tienen importantes implicaciones en la traducción entre lenguas afines (español-italiano), obteniendo cuatro grupos de desautomatizaciones: internas intencionales y no intencionales, y externas intencionales y no intencionales. Es evidente que los límites entre la idiomatidad y la literalidad no están siempre bien delimitados, de ahí que ciertas expresiones puedan resultar metafóricas para algunos y menos para otros. En efecto, cabe apuntar que la afinidad lingüística entre dos idiomas no es sinónimo de afinidad metafórica e idiomática (Trovato 2018, 426), de ahí que pueda resultar novedoso e interesante abordar la relación entre humor y fraseología desde el terreno de la lingüística computacional.

Gonzalo Moreno Cabrera en su estudio titulado «Los fraseologismos bíblicos en español, francés e inglés: estudio contrastivo y traductológico», selecciona un grupo conformado por 16 unidades fraseológicas extraídas del mundo bíblico y centra su interés en la vertiente traductológica, echando mano de la importante contribución que procede de la lexicografía bilingüe. Efectivamente, para la traducción de la fraseología el patrimonio cultural compartido entre las lenguas y sus culturas es de indudable relevancia, lo cual queda demostrado por los llamados *universales fraseológicos* o *européismos* (cf. Corpas Pastor 2000).

«Los marcos situacionales como estrategia para la traducción de enunciados fraseológicos del portugués y del español» de Ana Belén García Benito se propone mejorar la competencia fraseológica del traductor mediante la elaboración de repertorios bilingües de marcos situacionales organizados por situaciones prácticas y lingüísticas comunes a las dos lenguas. Esta estrategia se revela útil de cara a la actividad traductora en su sentido más amplio. Este estudio pone de relieve la importancia de los estudios de fraseología contrastiva y comparada en perspectiva lexicográfica, didáctica y traductológica, como señala, entre otros autores Greco (2020), en su estudio basado en la comparación interlingüística.

El segundo bloque está formado por cinco capítulos que profundizan en los aspectos contrastivos de la fraseología, mediante un acercamiento empírico y fundamentado en corpus. Da comienzo a la sección Paola Carrión González con su trabajo «Lingüística de corpus aplicada a la lexicografía haitiano-español: normalización y fraseología contrastiva». Lo novedoso de esta contribución reside en el abordaje de una lengua criolla, lo cual representa todo un desafío para el lingüista, especialmente en lo que a la dimensión fraseológica y a su tratamiento se refiere. De hecho, la autora afronta el diseño de cor-

pus como paso previo a la creación de una herramienta lexicográfica bilingüe para la combinación de lenguas haitiano-español. De esta manera, sería posible aportar una contribución al proceso de normalización de la lengua haitiana en el plano de la fraseología y contribuir, de paso, a la traducción interlingüística y cultural de la fraseología en ambos idiomas, como ponen de relieve Bassnett y Lefevre (1990), ahondando en la relación profunda que se establece entre la traducción y la cultura.

El capítulo a cargo de Beatriz Martín-Gascón titulado «El léxico en el discurso feminista y antifeminista en España y Estados Unidos: un estudio fraseológico y cognitivo sobre ‘feminista’ y ‘feminazi’ en Twitter», se basa en el Modelo Cognitivo Idealizado (MCI) y se propone abordar las metonimias, metáforas conceptuales y unidades fraseológicas relacionadas con los términos ‘feminista’, ‘feminist’ y ‘feminazi’, a partir de un corpus extraído de Twitter. Los resultados obtenidos demuestran que Twitter se perfila como un espacio socio-digital favorable al discurso de resistencia y la lucha feminista y apuntan a que el español y el inglés comparten una estructura conceptual y fraseológica similar.

«Un enfoque contrastivo del Sujeto + Animado en la construcción media inglesa y española» de Macarena Palma Gutiérrez se apoya en el marco de la Gramática de Construcciones de Goldberg y se propone establecer una serie de características comunes entre el inglés y el español a nivel semántico. Se analizan, en concreto, aquellas construcciones medias que no experimentan cambios en la valencia semántica y se pone de manifiesto que este fenómeno resulta más evidente con los verbos de emoción que implican una conceptualización a raíz de un estímulo externo.

El capítulo siete «Colocaciones verbales en el turismo de aventura: estudio contrastivo inglés-español» de Isabel Durán-Muñoz y Eva Lucía Jiménez-Navarro presenta los resultados preliminares de un proyecto de investigación en marcha en el que se lleva a cabo un estudio basado en corpus en lo referente a la fraseología de ámbito turístico, con especial referencia al turismo de aventura. El foco de interés son las colocaciones verbales que implican movimiento. A raíz del estudio, se han observado asimetrías en lo que atañe a los equivalentes de sus colocaciones y se ofrecen datos acerca de las estructuras argumentativas de los verbos analizados.

Cierra esta segunda sección el trabajo de María-Araceli Losey-León titulado «Estudio contrastivo de las colocaciones verbales en el ámbito marítimo en inglés y en español y representación fraseográfica con Oxygen XML Editor 22.1». Para este estudio que afronta terminología especializada (buque remolcador), la autora se vale de una técnica de extracción híbrida (estadística y lingüística) con el fin de analizar los datos, tanto desde un punto de vista semasiológico como onomasiológico. Su análisis concluye que en términos ge-

nerales existe una simetría en el plano sintáctico y también semántico y se pone de relieve, en última instancia, que la configuración actual de las nuevas tecnologías es emblemática y contribuye significativamente al tratamiento de la fraseología, como han señalado en un volumen de reciente publicación Corpas Pastor y Mitkov (2019).

La última sección, enfocada específicamente en cuestiones de traducción e interpretación en relación con la fraseología, se compone de cinco capítulos. El primero a cargo de María Angeles Recio y J. Agustín Torijano se titula «Errores de traducción por falsas equivalencias en las comparaciones enfáticas: el corpus como herramienta». En traducción, se considera el error como una desviación de la norma lingüística que puede afectar a la calidad del producto traducido (Hurtado Albir 2011, 289) y puede asumir varias formas en función del plano léxico, morfológico o sintáctico. Se analizan unidades fraseológicas formadas con verbos de emoción (temblar / llorar) y se realiza un análisis de los estereotipos subyacentes a las unidades seleccionadas para abordar el tema de los errores de traducción provocados por falsas equivalencias. Se reivindica, pues, la relevancia de la lingüística comparada y de la fraseología contrastiva para la didáctica de la traducción.

Melania Cabezas-García y Pilar León-Araúz abordan los términos poliléxicos como uno de los principales problemas de traducción en su capítulo titulado «Procedimiento para la traducción de términos poliléxicos con la ayuda de corpus». El concepto de problema de traducción está íntimamente relacionado con el de error (cf. Hurtado Albir 2011, 279) y remite a una dificultad objetiva a la que tiene que enfrentarse el traductor en su labor profesional. Ahora bien, los términos poliléxicos se configuran como locuciones terminológicas susceptibles de plantear muchas veces dificultades de comprensión y, por lo tanto, de traducción. Mediante técnicas heurísticas basadas en corpus paralelos y comparables, resulta posible agilizar la gestión traductológica de los términos poliléxicos por parte de los profesionales de la mediación lingüística.

«Colocaciones y locuciones en las fichas técnicas de impresoras 3D (inglés-español): un estudio basado en el corpus 3DCOR» de Ángela Luque-Giráldez y Míriam Seguiri se propone brindar al traductor un instrumento apto para atender sus necesidades de carácter terminológico al acometer una traducción de tipo técnico, o sea, un tipo de traducción especializada que se caracteriza por la univocidad y la monosemia (cf. Scarpa 2008). Para ello, las autoras han confeccionado un corpus bilingüe bidireccional virtual de fichas técnicas de impresoras 3D, garantizando su representatividad, tanto cualitativa como cuantitativa. De hecho, su análisis basado en corpus confirma la elevada tendencia a la sustantivación en el género de ficha técnica.

Carlos Manuel Hidalgo-Ternero coloca su trabajo, titulado «El algoritmo ReGap para la mejora de la traducción automática neuronal

de expresiones pluriverbales discontinuas (FR>EN/ES)», en el marco de la lingüística computacional. Para comprobar la eficacia del mencionado algoritmo, el autor lleva a cabo una experimentación con el sistema TAN más sólido hasta hoy conocido, esto es, Deepl. A este fin, aborda la traducción de somatismos verbonominales en las direcciones FR>EN y FR>ES. Los resultados recabados son prometedores e instan a seguir desarrollando el sistema de detección y conversión automáticas de ReGap con el propósito de probar su eficacia con otras parejas de lenguas, así como con otros tipos de expresiones pluriverbales.

El volumen termina con la aportación de Gloria Corpas Pastor y Mahmoud Gaber titulada «Extracción de fraseología para intérpretes a partir de corpus comparables compilados mediante reconocimiento automático del habla». Fundamentado en las modernas tecnologías del habla, este capítulo propone una innovadora metodología de compilación automatizada de corpus comparables de discursos orales y documentos escritos de cara a la preparación de un encargo de interpretación sobre cambio climático. Este trabajo brinda información sumamente relevante acerca de los rasgos diferenciales de la lengua hablada y abre nuevos e interesantes horizontes para los estudios de interpretación. Los resultados obtenidos apuntan al fenómeno de la simplificación en función de la complejidad del patrón colocacional y del término base, de ahí que a mayor complejidad, le corresponda un número inferior de colocados verbales tipo. En definitiva, los corpus orales aportan al intérprete más ventajas que los corpus escritos, particularmente en la fase de preparación y arrojan luz importante sobre la cuestión de las colocaciones.

3

Ahora bien, es bien sabido que la lingüística de corpus basa sus investigaciones en datos obtenidos a partir de corpus, esto es, muestras reales de uso de la lengua. Se trata, pues, de un enfoque metodológico que es posible adoptar desde disciplinas diversas, que se contraponen a una metodología basada en la introspección. Lo que acabamos de apuntar corrobora la relevancia científica de su uso en los análisis enfocados en la fraseología. Si a ello le añadimos los novedosos enfoques computacionales, podemos observar de qué manera se están consolidando las tendencias más actuales en el análisis del fenómeno fraseológico aplicado a la traductología, como señalamos en un trabajo reciente (Trovato 2021, 284). De hecho, el empleo de recursos informáticos y, más concretamente, de las TIC (Tecnologías de la información y la comunicación) para clasificar datos lingüísticos es ya una realidad que ha venido para quedarse. Su presencia, de ninguna manera, pretende y puede reemplazar a lingüistas y fra-

seólogos, sino que tiene la función de auxiliarlos en sus actividades de sistematización. Se trata, por lo tanto, de complementar y, sobre todo, de agilizar su trabajo. El hecho de poder consultar un corpus mediante un simple clic es una operación que hubiese resultado imposible hace dos décadas. Hoy, en cambio, es una acción a la orden del día como viene confirmando el volumen que hemos presentado en esta contribución. Este manual se presenta como una obra con carácter polifónico y cuenta con una dimensión multilingüe, ya que contiene estudios sobre alemán, español, francés, haitiano, inglés, italiano y portugués en distintas combinaciones y tiene como hilo conductor el planteamiento contrastivo y traductológico. El acto de traducción lleva inevitablemente aparejada la noción de comparación entre las dos lenguas involucradas en el proceso traslativo; una comparación que muchas veces asume la forma de contraste, especialmente cuando surgen diferencias en el plano formal. Sin embargo, es precisamente la actividad traductora la que permite poner en contacto a dos lenguas y dos culturas en su calidad de agente codificador y descodificador de la comunicación (Torop 2010, 231). Pues bien, si consideramos la lengua y la cultura como dos caras de la misma moneda, cabrá interrogarnos acerca de los mecanismos más pertinentes para trasladar los elementos culturales y metafóricos de una lengua A a una lengua B. A este respecto, Diadori (2018, 224-5) observa que la vertiente cultural y metafórica representa la manifestación verbal más evidente de una lengua y, por esta razón, puede dificultar la tarea traductora. Lo que acabamos de apuntar pone de manifiesto la estrecha relación entre la traducción y la fraseología, hoy en día auxiliadas por las nuevas y modernas tecnologías informáticas.

Para terminar, los lectores potenciales del volumen aquí presentado encontrarán en él no solo valiosos motivos de reflexión sobre el fenómeno abordado, sino también útiles y eficaces aplicaciones para llevar al aula y para implementar en traducción e interpretación concebidas como prácticas profesionales.

Bibliografía

- Bassnett, S.; Lefevere, A. (eds) (1990). *Translation, History and Culture*. London; New York: Pinter.
- Casares, J. (1992). *Introducción a la lexicografía moderna*. Madrid: Consejo Superior de Investigaciones Científicas.
- Corpas Pastor, G. (1996). *Manual de fraseología española*. Madrid: Gredos.
- Corpas Pastor, G. (2000). «Acerca de la (in)traducibilidad de la fraseología». Copas Pastor, G. (ed.), *Las lenguas de Europa: estudios de fraseología, fraseografía y traducción*. Granada: Editorial Comares, 483-522.
- Corpas Pastor, G. (2003). *Diez años de investigación en fraseología: análisis sintáctico-semánticos, contrastivos y traductológicos*. Madrid: Iberoamericana.
- Corpas Pastor, G.; Bautista Zambrana, M.R.; Hidalgo-Terreno, C.M. (2021). *Sistemas fraseológicos en contraste: enfoques computacionales y de corpus*. Granada: Editorial Comares.
- Corpas Pastor, G.; Colson, J.P. (eds) (2020). *Computational Phraseology*. Amsterdam: John Benjamins. IVITRA Research in Linguistics and Literature | Studies, Editions and Translations 24.
- Corpas Pastor, G.; Mitkov, R. (eds) (2019). *Computational and Corpus-Based Phraseology = Proceedings of the Third International Conference, Europhras 2019* (Malaga, Spain, 25-27 September 2019). Berlin: Springer.
- Diadori, P. (2018). *Tradurre: una prospettiva interculturale*. Roma: Carocci Editore.
- Greco, S. (2020). «Combinaciones léxicas fijas y determinante en español e italiano». *Artifara*, 20, 127-49. <https://doi.org/10.13135/1594-378X/3985>.
- Hurtado Albir, A. (2011). *Traducción y Traductología. Introducción a la Traductología*. Madrid: Cátedra.
- Martínez Marín, J. (1996). *Estudios de fraseología española*. Málaga: Ágora.
- Mitkov, R.; Monti, J.; Copras Pastor, G.; Seretan, V. (eds) (2018). *Multiwords Units in Machine Translation and Translation Technology*. Amsterdam; Philadelphia: John Benjamins. Current Issues in Linguistic Theory 341.
- Ruiz Gurillo, L. (1998). *La fraseología del español coloquial*. Barcelona: Editorial Ariel SA.
- Scarpa, F. (2008). *La traduzione specializzata. Un approccio didattico professionale*. Milano: Hoepli.
- Torop, P. (2010). *La traduzione totale*. Milano: Hoepli.
- Tristá Pérez, A.M. (1988). *Fraseología y contexto*. La Habana: Editorial de Ciencias Sociales.
- Trovato, G. (2018). «Un estudio contrastivo y pedagógico español-italiano acerca del tratamiento lingüístico-traductológico de unidades léxicas dotadas de carácter figurado e idiomático». *Orillas. Rivista d'Ispanistica*, 7, 417-28. http://orillas.cab.unipd.it/orillas/07_09trovato_astilleros/.
- Trovato, G. (2021). «Nuevas perspectivas de investigación en la traducción especializada en lenguas románicas: aspectos comparativos, léxicos, fraseológicos, discursivos y didácticos». *Mutatis Mutandis. Revista Latinoamericana de Traducción*, 14, 277-84. <https://doi.org/10.17533/udea.mut.v14n2a01>.
- Zuluaga Ospina, A. (1980). *Introducción al estudio de las expresiones fijas*. Frankfurt a.M.: Lang