

PAREMIOLOGÍA BASADA EN CORPUS WACKY: ENFOQUE (INTRA- E INTER-) LINGÜÍSTICO Y CONCEPTUAL

Vincenzo Lambertini

Università di Bologna

Dipartimento di Interpretazione e Traduzione,
sede di Forlì

vincenzo.lambertini2@unibo.it

Resumen

Este estudio, que forma parte de un proyecto de investigación doctoral llevado a cabo en el Departamento de Interpretación y Traducción (DIT) de Forlì (Universidad de Bologna), se propone categorizar de manera conceptual los refranes italianos de uso común, intentando facilitar una herramienta de gran utilidad para intérpretes, traductores y hablantes que quieran expresar determinados conceptos utilizando refranes.

Los obstáculos encontrados en esta investigación han sido muchos, ya que los estudios lingüísticos sobre paremias escasean.

No obstante, gracias al corpus italiano itWaC, se pudo hallar una metodología de identificación automática de paremias, además de llevar a cabo un estudio lingüístico sobre refranes y sus modificaciones.

Las etapas futuras consistirán en realizar una categorización conceptual de los refranes y aplicar los mismos principios de identificación automática de paremias y análisis conceptual a refranes de otros idiomas, para realizar un recurso multilingüe, utilizable de manera semasiológica para recuperar refranes en diferentes lenguas.

1. INTRODUCCIÓN

Este artículo está encaminado a describir la primera fase de un estudio de investigación que se propone categorizar de manera conceptual los refranes italianos de uso común, con el fin de facilitar una herramienta de gran utilidad para intérpretes, traductores y hablantes que quieran expresar determinados conceptos utilizando refranes.

Para darse cuenta de las dificultades relacionadas con el estudio de los refranes y su categorización, es necesario dar un paso atrás y centrarse en aspectos teóricos de suma importancia.

En primer lugar, cabe destacar la cuestión de la fijación de los refranes: según numerosos investigadores, entre los cuales hemos de recordar a Jean-Claude Anscombe, los refranes “no son expresiones fijas ni tampoco giros idiomáticos” (Anscombe, 1997, p.52).

En segundo lugar, la falta de fijación implica, en la mayoría de los casos, que el sentido de un enunciado es composicional, es decir se deduce del significado de sus partes y de las reglas con las que se combinan (Casadei, 1996; véase también Katz, 1973). Es evidente que los refranes no son frases idiomáticas, pero tampoco son frases composicionales ni siquiera

funcionan semánticamente como secuencias libres. Es suficiente poner un ejemplo de correspondencia interlingüística de refranes (español, francés e italiano) para demostrar su funcionamiento no composicional.

(1a) Hablando del rey de Roma por la puerta asoma

(1b) *Quand on parle du loup on en voit la queue* (trad. lit.: “Cuando se habla del lobo se le ve la cola”)

(1c) *Si parla del diavolo e spuntano le corna* (trad. lit.: “Se habla del diablo y aparecen los cuernos”)

Como se puede observar, los tres refranes tienen sentidos literales muy distintos. No obstante, al considerar sus definiciones, se aprecia que los tres refranes tienen el mismo sentido, puesto que apuntan a la misma situación genérica, que podríamos resumir citando la explicación del Refranero Multilingüe (2015): “se emplea cuando se presenta inesperadamente la persona de la que se está hablando”.

Como se deriva de lo argumentado hasta aquí, partimos de la idea de que los refranes tienen cierto grado de fijación y no significan de inmediato lo que pretenden comunicar composicionalmente. De hecho, transmiten un mensaje metafórico, y más precisamente genérico, que tiene rasgos invariados en distintos contextos.

En el presente artículo, se describen en primer lugar las dificultades con las que nos encontramos al intentar identificar automáticamente refranes. A continuación, se explica el modo cómo logramos identificar refranes italianos y sus versiones modificadas. Por último, se exponen brevemente las conclusiones a las que se ha llegado por el momento y los proyectos de investigación futuros.

2. IDENTIFICACIÓN AUTOMÁTICA DE REFRANES

En muchas ocasiones, los estudios sobre refranes prescinden de datos reales, descuidando las características lingüísticas (semánticas, sintácticas y morfológicas) y concentrándose más en lo folklórico, histórico y cultural. El límite de estos planteamientos es que no consideran los refranes elementos empleados en el habla contemporánea, mientras que en muchos ámbitos, como en el lenguaje publicitario o periodístico, parecen imprescindibles. Por esta razón, es preciso analizar con detenimiento los refranes en corpus lingüísticos.

No cabe duda de que no es sencillo encontrar automáticamente refranes en corpus de grandes dimensiones, pero tampoco es imposible. En nuestro trabajo, decidimos tomar como referencia las investigaciones sobre metáforas³⁷, pues comparten diversos aspectos con los refranes.

En principio, hay varios métodos para identificar automáticamente metáforas en corpus. Por ejemplo, se puede analizar manualmente un corpus y decidir si una palabra se utiliza de manera metafórica. Sin embargo, este método no se puede llevar a cabo con corpus de grandes entidades, imprescindibles para estudiar refranes, ya que su frecuencia de uso no es muy elevada. Otro sistema podría consistir en buscar palabras o unidades fraseológicas utilizando lemas precisos. El límite de este método es que el investigador interfiere demasiado en la búsqueda y no permite al corpus producir autónomamente los resultados,

³⁷ Por ejemplo, véase Deignan (2010).

lo que en cambio debería ocurrir de acuerdo con los principios clave de la lingüística de corpus (Tognini-Bonelli, 2001).

Para intentar no influir en la búsqueda de refranes, decidimos tomar una lista de más de 25.000 refranes italianos sacados del *Dizionario dei proverbi italiani* de Carlo Lapucci (2006), una de las obras más completas sobre el acervo paremiológico italiano, que además numera cada uno de los refranes. Generamos, pues, una lista de 500 números casuales y tomamos en consideración 500 refranes correspondientes a esos números casuales. Al final, descubrimos que solo un 4,16% de los 500 refranes detectados casualmente estaba presente en nuestro corpus, lo que indicó que esta metodología no podía compaginarse con nuestro estudio.

Según los estudios sobre metáforas, otra manera de encontrar metáforas en corpus consiste en la búsqueda de los llamados “marcadores de metáforas”, a saber aquellas expresiones que introducen una metáfora (Deignan, 2010, p.17). Se trata de locuciones como “una suerte de”, “por así decirlo”, etc. Este método no garantiza la recuperación automática de todas y cada una de las metáforas presentes en un corpus, pero sí permite identificar metáforas sin que el investigador interfiera en este proceso.

Es posible aplicar estos principios al estudio sobre refranes, dado que, como subraya Shapira (2000, pp.89-90), las paremias pueden ser introducidas por expresiones que señalan su carácter paremiológico. Sin embargo, una vez más, los estudios lingüísticos sobre refranes en italiano escasean, así que no hay datos ciertos sobre los “marcadores de refranes” italianos más frecuentes. Es más, al poner en práctica estos principios se notó que con expresiones introductorias muy generales (por ejemplo: “*come si suol dire*”, “*come si dice*”, etc.) no solo se obtienen refranes, sino también otros tipos de fraseologismos y frases sentenciosas, como expresiones idiomáticas, máximas, aforismas, dichos, etc.

Para superar esta dificultad, se decidió buscar la palabra “*proverbio*”³⁸ que está contenida con una probabilidad elevada en expresiones como: “*come dice il proverbio*”; “*c’è un proverbio che dice*”; “*come recita il proverbio*”; etc. Se trata, pues, de la única parte que está presente con elevada frecuencia y probabilidad en los marcadores de refranes.

3. TIPOLOGÍA DEL CORPUS

El corpus itWaC (Baroni, et al., 2008) forma parte de una serie de corpus de varios idiomas realizados a partir de textos recuperados en la Web reunidos bajo la etiqueta “Corpus WaCky”, lo que es acrónimo de “*The Web-As-Corpus Kool Yinitiative*”. El corpus italiano itWaC se compone de dos mil millones de palabras y está formado por textos recuperados en la Web con dominio .it, utilizando como *seeds* las palabras de frecuencia mediana del corpus “*La repubblica*” y un vocabulario básico de lengua italiana.

Además de ser un corpus actual que abarca varias tipologías de textos, que van desde los foros hasta los periódicos electrónicos, este corpus está etiquetado con información morfosintáctica (o *POS tagging*), lo que resultará muy útil a la hora de buscar refranes, puesto que, como se verá más adelante, estos pueden variar mucho.

Finalmente, este corpus puede ser consultado gracias al programa gratuito NoSketch Engine que está disponible en la dirección: http://nl.ijs.si/noske/wacs.cgi/first_form.

³⁸ “*Proverbio*” corresponde al español “refrán” y no a “proverbio”. La diferencia entre el “*proverbio*” italiano y el “proverbio” español es que el segundo indica un tipo de paremia culta (Sevilla Muñoz, 2000, p.101), mientras que el primero no tiene un sentido tan restrictivo.

Además, es posible abrir el corpus itWaC directamente desde esta página, seleccionándolo en el menú desplegable “Corpus”.

4. METODOLOGÍA DE BÚSQUEDA DE REFRANES EN ITWAC

Al insertar la palabra “*proverbio*” en la interfaz de consulta del corpus itWaC, se obtienen 3706 resultados que corresponden a las concordancias de la palabra “*proverbio*”. Esto significa que en el contexto de cada uno de estos ejemplos se habla de refranes y muy probablemente están presentes refranes. Sin embargo, no es fácil validar de manera electrónica los resultados, ya que el procedimiento que tendría que realizar la máquina sería demasiado complejo. Por esta razón, fue necesario validar cada uno de los resultados manualmente, para distinguir entre refranes utilizados en contexto y refranes de los que se habla sin que estén presentes, además de otros casos que no son relevantes para este estudio.

En la fase de validación, se agruparon los resultados obtenidos en seis clases, dependiendo de la presencia / ausencia de los refranes y también de su tipología. De este modo, se observó que el 92% de los resultados analizados contenía refranes y el 8% restante no presentaba ningún refrán. Es más, notamos que un 64% consistía en refranes utilizados en contexto, dividido en un 28% de refranes italianos y un 36% de refranes no italianos (citados en lengua original o traducidos literalmente al italiano), lo que es muy llamativo y necesitaría ser objeto de estudios más específicos. El 8% de los resultados se refería a refranes pertenecientes a obras literarias italianas anteriores al siglo XX y que ya son en desuso, y otro 7% simplemente presentaba reflexiones sobre refranes y costumbres del pasado. Finalmente, un 13% estaba compuesto por repeticiones generadas por el programa de búsqueda.

Tras centrarnos en los refranes italianos, pudimos detectar 690 refranes. Algunos de estos ocurren más de una vez, lo que queda demostrado por el hecho de que en total los refranes italianos hallados son 1037. Esto no quiere decir que se repitan los mismos resultados, sino que se utilizan los mismos refranes en diferentes textos. Por esto, decidimos realizar una lista de todos los refranes que identificamos con las respectivas ocurrencias para lograr dos objetivos principales. El primero consistía en buscar los refranes encontrados en el propio corpus, para averiguar si estos se utilizaban también en otras ocasiones donde no aparecía la palabra “*proverbio*”. El segundo objetivo estaba encaminado a detectar la frecuencia de uso de los refranes hallados, además de su contexto, para proponer una lista de frecuencia de uso de los refranes italianos que aún no habíamos encontrado.

También en este caso los resultados son muy llamativos. Por lo que respecta a los refranes detectados al principio utilizando la palabra “*proverbio*”, se nota que una mayoría aplastante (el 77%) tiene una sola ocurrencia. El 11% tiene 2 ocurrencias, el 5% tiene 3, el 4% tiene 4, un 1% tiene 5 ocurrencias, otro 1% tiene 6 y finalmente otro 1% tiene entre 8 y 7 ocurrencias.

Tras obtener la lista de los 690 refranes, fue posible buscar los mismos refranes en el corpus para averiguar su funcionamiento en el contexto. Este estudio se limitó únicamente a los refranes de ocurrencias entre 8 y 2, puesto que, como ya se ha subrayado en este artículo, el objetivo principal ha sido encontrar una metodología de investigación que se pueda aplicar a otros tipos de estudios en este ámbito. De hecho, una vez determinadas las pautas adecuadas para llevar a cabo investigaciones parecidas sobre refranes, solo haría falta

aplicarlas a refranes encontrados en corpus distintos y también a refranes hallados en corpus de otros idiomas (con los ajustes necesarios para respetar las características de las diferentes lenguas).

Antes de presentar los resultados detallados de esta consulta, sería oportuno hacer hincapié en datos estadísticos sobre los mismos. Como sugiere Shapira (2000, pp.89-92) sobre el uso de refranes en lengua francesa, la mayoría de las veces los refranes no se introducen con un marcador de refrán, sino que se citan directamente. Los números que atañen a los refranes más frecuentes hablan muy claro: el refrán que más se utiliza en el corpus de referencia es “*L’union fa la forza*” (trad. lit.: “La unión hace la fuerza”). De hecho ocurre en su forma estándar y en diversas versiones modificadas 579 veces, mientras que estaba presente en la lista de refranes obtenida utilizando la palabra “*proverbio*” tan solo 5 veces. El segundo refrán más frecuente es “*Il buongiorno si vede dal mattino*” (trad. lit.: “El buen día se ve desde la mañana”), que aparece en el corpus 447 veces en su forma estándar o en versiones modificadas, y que tenía 4 ocurrencias al buscarlos utilizando la palabra “*proverbio*”. Lo mismo puede decirse sobre el refrán “*Non c’è fumo senza arrosto*” (trad. lit.: “No hay humo sin asado”), que tenía 4 ocurrencias y que, en cambio, se repite en total 382 veces.

Como ya se ha adelantado, la mayoría de los refranes se inserta directamente en el discurso sin ninguna fórmula introductoria. Sin embargo, otras observaciones subrayan su flexibilidad de uso. De hecho, una parte importante de la literatura paremiológica indica que generalmente las paremias se modifican. Pensemos por ejemplo en el estudio de Shapira (2000) que habla de *déproverbialisation* y *détournement* o, en español, desautomatización (García Yelo, 2012). Por consiguiente, nuestra atención se centró en las modificaciones de los refranes, para encontrar en el corpus no solo los refranes en su forma estándar³⁹, sino también sus diferentes variantes y sus versiones modificadas.

El corpus, con su clasificación y etiquetado por partes del discurso (o *POS tagging*), posibilita encontrar secuencias buscando tanto lemas específicos como partes del discurso, lo que permite hallar refranes modificados. Pero antes establecimos una serie de pautas para que la búsqueda de refranes modificados fuera igual y comparable para cualquier paremia de nuestra lista. Una vez más, hemos de subrayar que no existen estudios específicos sobre cómo se modifican los refranes y qué se puede modificar en los refranes.

Por esto, la primera regla que nos dimos fue tomar como referencia las modificaciones y variantes de los refranes que encontramos en el propio corpus gracias a la búsqueda de la palabra “*proverbio*”. Por ejemplo, el refrán “*Chi ben comincia è a metà dell’opera*” (trad. lit.: “Quien bien empieza está a mitad de la obra”) es la única variante aceptada por *Il grande dizionario dei proverbi italiani* (2006). No obstante, en el corpus analizado se notan otras variantes del mismo refrán, como “*Chi ben **inizia** è **già** a metà dell’opera*”; “*Chi ben comincia è **alla** metà dell’opera*”; “*Chi ben **principia** è **alla** metà dell’opera*”; “*Chi ben **incomincia** è a metà dell’opera*” (en negrita las palabras diferentes con respecto a la variante aceptada por el diccionario citado). Otro refrán que no siempre respeta las indicaciones del diccionario es “*Se la montagna non va da Maometto, Maometto va alla montagna*”⁴⁰ (trad. lit.: “Si la montaña no

³⁹ En este estudio, no se entiende por “forma estándar” la variante presentada por los diccionarios u otros recursos dedicados a los refranes, sino la variante encontrada en el corpus con la primera búsqueda, es decir buscando la palabra “*proverbio*”.

⁴⁰ Esta es la variante aceptada por otro diccionario, el de Lapucci (2006), ya que este refrán no está presente en *Il grande dizionario dei proverbi italiani* (2006). Esto es muy llamativo, puesto que se trata de uno de los refranes más frecuentes en nuestro corpus, lo que indica su elevada frecuencia de uso en el italiano contemporáneo. Se podría pensar que los recursos que deberían dar cuenta de la presencia y uso de los refranes, además de su sentido, no siempre tiene la prioridad de describir del habla corriente sino más bien recopilar refranes, sin hacer distinción entre los realmente empleados y los anticuados.

va a Maoma, Maoma va a la montaña”). En nuestro corpus, los autores de los textos donde aparece dicho refrán se confunden (no se sabe si de manera voluntaria o involuntaria) y escriben: “*Se Maometto non va alla montagna la montagna va da Maometto*” (la coma entre las dos oraciones ya no está presente en los textos originales). En el corpus, se hallaron otros ejemplos de modificaciones: “*Se Maometto non va alla montagna, è la montagna che va da Maometto*” y “*Se la montagna non va da Maometto, è Maometto **ad andare** alla montagna*”.

El análisis de las modificaciones ya presentes en un corpus lleva al investigador a intentar averiguar cómo se modifican los refranes y cuáles son los elementos que pueden variar. El problema surge para los refranes que solo aparecen una vez al buscarlos utilizando los marcadores de refranes. Por esa razón, intentamos detectar una metodología de búsqueda de refranes modificados basada en los siguientes principios:

En primer lugar, para cada uno de los refranes se insertó, en la interfaz de consulta del corpus, solo la primera parte y, posteriormente, solo la segunda. Para decidir cómo dividir los refranes nos inspiramos en los estudios sobre refranes bimembres (Honeck, 1997, pp. 51-57). Si los refranes estudiados no eran bimembres, intentamos insertar solo las dos primeras palabras y, si no obtuvimos resultados satisfactorios, fuimos introduciendo una palabra más de cada vez, hasta encontrar resultados significativos. Esto sirve en general para encontrar refranes citados solo parcialmente, lo que es muy frecuente y forma parte de lo que Shapira (2000) define *déproverbialisation*.

En segundo lugar, se decidió sustituir las palabras de los refranes por etiquetas morfosintácticas de esta manera: el objetivo, y a la vez el método para seguir, era encontrar la combinación entre palabras y etiquetas morfosintácticas que permitiera hallar más refranes que enunciados libres utilizando el menor número de lemas y el mayor número de etiquetas morfosintácticas. Se tome por ejemplo el refrán (2a).

(2a) *Una mela al giorno toglie il medico di torno* (trad. lit.: “Una manzana cada día quita al médico de encima”)

Si se utiliza una secuencia de búsqueda formada enteramente por etiquetas morfosintácticas como (2b), se encuentran refranes, pero también una gran cantidad de enunciados que no tienen nada que ver con (2a).

(2b) [artículo] + [sustantivo] + [preposición compuesta] + [sustantivo] + [verbo] + [artículo] + [sustantivo] + [preposición simple] + [sustantivo]

Si se empieza a sustituir las etiquetas morfosintácticas con los lemas presentes en el refrán, se nota que los lemas necesarios son “*al*”, “*giorno*”, “*di*”, que podrían incluso constituir la estructura más fija e invariable del refrán. Asimismo, se observa que la combinación constituida por el menor número de palabras y el mayor número de etiquetas morfosintácticas, que produce resultados relevantes, es (2c).

(2c) [artículo] + [sustantivo] + “*al*” + “*giorno*” + [verbo] + [artículo] + [sustantivo] + “*di*”

Como se ha notado, el hecho de encontrar refranes gracias a su estructura mínima invariable podría ser muy útil a la hora de producir nuevas reflexiones sobre qué pieza léxica se puede modificar en los refranes y cuál no se debe modificar.

5. CONCLUSIONES

Después de una introducción general sobre los asuntos más relevantes en el terreno de la lingüística de los refranes, aún poco estudiados desde el punto de vista lingüístico, la

investigación que nos está ocupando se propone realizar un estudio lingüístico sobre refranes, a través de la observación de datos reales sacados de itWaC, un corpus de lengua general italiana de dos mil millones de palabras.

El primer avance de este estudio consiste precisamente en llevar a cabo una investigación paremiológica enteramente basada en datos reales: los refranes han sido detectados de forma casi automática, buscando lo que hemos definido “marcadores de refranes”. En realidad, el único elemento que desarrolla este papel y es capaz de producir la cantidad más elevada de refranes, junto con el porcentaje menor de resultados no relevantes, es la palabra italiana “*proverbio*” (refrán). De todos modos, sería importante estudiar todas las expresiones que suelen acompañar a los refranes, para dar pasos adelante en la identificación automática de paremias.

Por otra parte, gracias al etiquetado morfosintáctico del corpus, no solo es posible elaborar una lista de frecuencia de refranes, sino también buscar refranes en su variante estándar y en sus versiones modificadas.

Los datos demuestran que los refranes, si bien pueden modificarse, se cimientan en estructuras básicas y fijas, que permiten identificar la forma estándar de las paremias modificadas. Una investigación de este tipo lleva asimismo a detectar estas estructuras estables, desde un punto de vista tanto léxico como sintáctico. La etapa siguiente, pues, será buscar estructuras fijas desde la óptica semántica, lo que permitiría categorizar los refranes según su sentido “idiomático” y no “literal”. Todo esto posibilitaría la búsqueda y recuperación de refranes categorizados conceptualmente en distintos idiomas, superando los límites de los recursos actuales en el ámbito paremiológico. En otros términos, se podría pasar desde la perspectiva actual de tipo semasiológico hasta una perspectiva futura de carácter onomasiológico.

Aún queda mucho por investigar, pero estamos seguros de que el camino emprendido es capaz de llevar a resultados novedosos no solo en el marco de los refranes sino también en otras disciplinas, como la fraseología, la lexicografía y la terminología.

Bibliografía

- ANSCOMBRE, J.-C., 1997. Reflexiones críticas sobre la naturaleza y el funcionamiento de las paremias. *Paremia*, 6, pp.43-54.
- ANSCOMBRE, J.-C., 2005. Les proverbes : un figement du deuxième type ? *Linx*, 53, [online] Available at: <http://linx.revues.org/255> [Accessed 11 October 2012].
- BARONI, M., BERNARDINI, S., FERRARESI, A. AND ZANCHETTA, E., 2009. The WaCky Wide Web: A Collection of Very Large Linguistically Processed Web-Crawled Corpora. *Language Resources and Evaluation*, 43(3), [online] Available at: http://wacky.sslmit.unibo.it/lib/exe/fetch.php?media=papers:wacky_2008.pdf [Accessed 9 October 2014].
- BARSANTI VIGO, M.J., 2006. Problemática en torno al refrán y otras categorías paremiales. In: M. Alonso Ramos, ed. 2006. *Diccionarios y fraseología. Anexos de Revista de Lexicografía*, 3, pp.197-206.
- BIDAUD, F., 2002. *Structures figées de la conversation. Analyse contrastive français-italien*. Bern: Lang.
- BOGGIONE, V. AND MASSOBRIO, L., 2007. *Dizionario dei proverbi*. Torino: UTET.
- CAMPOS, J. G. AND BARELLA, A., 1995. *Diccionario de refranes*. Madrid: Espasa Calpe.

- CANELLADA, M. J. AND PALLARES, B., 2001. *Refranero español. Refranes, clasificación, significación y uso*. Madrid: Castalia.
- CASADEI, F., 1996. *Metafore ed espressioni idiomatiche: uno studio semantico sull'italiano*. Roma: Bulzoni.
- CORREAS, G., 1992. *Vocabulario de refranes y frases proverbiales y otras fórmulas comunes de la lengua castellana en que van todos los impresos antes y otra gran copa que juntó el Maestro Gonzalo Correas*. Madrid: Visor.
- CRISTILLI, C., 1989. Il proverbio come esempio di testualità popolare. In: C. Vallini, ed. 1989. *La pratica e la grammatica. Viaggio nella linguistica del proverbio*. Napoli: Istituto Universitario Orientale, Dipartimento di studi letterari e linguistici dell'Occidente. pp.177-206.
- DEIGNAN, A., 2009. Searching for Metaphorical Patterns in Corpora. In: P. Baker, ed. 2009. *Contemporary Corpus Linguistics*. London, New York: Continuum. pp.9-31.
- DEPECKER, L., 2011. Comment aborder le concept d'un point de vue linguistique ? In : J.-J. Briu, 2011. *Terminologie (I) : analyser des termes et des concepts*. Bern: Lang. pp.17-32.
- DUNETON, C. AND CLAVAL, S., 1990. *Le bouquet des expressions imagées : encyclopédie thématique des locutions figurées de la langue française*. Paris: Seuil.
- GARCÍA YELO, M., 2012. El proceso de desautomatización de paremias españolas en las redes sociales. *Unidades fraseológicas y TIC*. [online] Serie «Monografías», n°2, pp.111-124. Available at: http://cvc.cervantes.es/lengua/biblioteca_fraseologica/n2_gonzalez/default.htm [Accessed 20 August 2013].
- GONZÁLEZ REY, M.I. ed., 2012. *Unidades fraseológicas y TIC*. [online] Madrid: Instituto Cervantes, Biblioteca fraseológica y paremiológica, Serie «Monografías», n°2. Available at: http://cvc.cervantes.es/lengua/biblioteca_fraseologica/n2_gonzalez/default.htm [Accessed 20 August 2013].
- GROSS, G., 1996. *Les expressions figées en français*. Gap: Ophrys.
- GUZZOTTI, P. AND ODDERA, M.F., 2010. *Il grande dizionario dei proverbi italiani*. Bologna: Zanichelli.
- HERBST, T., FAULHABER, S. AND UHRIG, P. eds., 2011. *The Phraseological View of Language. A Tribute to John Sinclair*. Berlin: De Gruyter Mouton.
- HONECK, R.P., 1997. *A proverb in mind: the cognitive science of proverbial wit and wisdom*. Mahwah (N. J.): Lawrence Erlbaum associates.
- JUNCEDA, L., 1991. *Del dicho al hecho*. Barcelona: Obelisco.
- KATZ, J.J., 1973. Compositionality, idiomaticity and lexical substitution. In: S. R. Anderson, P. Kiparsky, eds. 1973. *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston. pp.357-376.
- KLEIBER, G., 1990. *La sémantique du prototype. Catégories et sens lexical*. Paris: PUF.
- KLEIBER, G., 1999. Les proverbes : des dénominations d'un type « très très spécial ». *Langue française*, 123, pp.52-69.
- KLEIBER, G., 2000. Sur le sens des proverbes. *Langages*, 139, pp.39-58.

- LAKOFF, G. AND JOHNSON, M., 1980. *Metaphors We Live By*. Chicago: University of Chicago.
- LAPUCCI, C., 1993. *Dizionario dei modi di dire della lingua italiana*. Milano: Garzanti.
- NORRICK, N.R., 1985. *How proverbs mean: semantic studies in English proverbs*. Berlin: Mouton.
- PERRIN, L., 2000. Remarques sur la dimension générique et sur la dimension dénomminative des proverbes. *Langages*, 139, pp.69-80.
- PHILIP, G., 2011. *Colouring Meaning. Collocation and connotation in figurative language*. Amsterdam: John Benjamins.
- PITTANO, G., 1996. *Frase fatta capo ha: dizionario dei modi di dire, proverbi e locuzioni*. Bologna: Zanichelli.
- QUARTU, B.M., 2001. *Dizionario dei modi di dire della lingua italiana: 10.000 modi di dire ed estensioni figurate in ordine alfabetico per lemmi portanti e campi di significato*. Milano: Rizzoli.
- QUIROGA, P., 2006. *Fraseología italo-española. Aspectos de lingüística aplicada y contrastiva*. Granada: Granada lingüística.
- SEVILLA MUÑOZ, J. AND ZURDO RUIZ-AYÚCAR, M.I.T. eds., 2009. *Refranero multilingüe*. Madrid. Instituto Cervantes (Centro Virtual Cervantes). [online] Available at: **¡Error! Referencia de hipervínculo no válida.** [Accessed 20 March 2015].
- SEVILLA MUÑOZ, J., 2000. Les proverbes et phrases proverbiales français, et leurs équivalences en espagnol. *Langages*, 139, pp.98-109.
- SHAPIRA, C., 2000. Proverbe, proverbialisation et déproverbialisation. *Langages*, 139, pp.81-97.
- SIMPSON, J., 1992. *The Concise Oxford Dictionary of Proverbs*. Oxford: Oxford University Press.
- SINCLAIR, J., 2003. *Reading Concordances. An Introduction*. London: Pearson Longman.
- TOGNINI-BONELLI, E., 2001. *Corpus Linguistics at Work*. Amsterdam: John Benjamin.
- TURRINI, G., ALBERTI, C., SANTULLO M.L. AND ZANCHI, G., 1995. *Capire l'antifona*. Bologna: Zanichelli.
- VISETTI, Y.M. AND CADIOT, P., 2006. *Motifs et proverbes. Essai de sémantique proverbiale*. Paris: Presses Universitaires de France.
- XATARA, C.M., 2009. *Dictionnaire d'expressions idiomatiques: Français – Portugais – Français*. [online] Available at: http://www.cnrtl.fr/dictionnaires/expressions_idiomatiques/ [Accessed 10 March 2015].