UNIVERSIDAD DE MÁLAGA

PROGRAMA DE DOCTORADO EN
LINGÜÍSTICA, LITERATURA Y TRADUCCIÓN

Facultad de Filosofía y Letras
Departamento Traducción e Interpretación
Universidad de Málaga

Tesis doctoral

INDICADORES FRASEODIDÁCTICOS: UNA PROPUESTA DE SISTEMATIZACIÓN DE LA FRASEOLOGÍA ITALIANA SEGÚN EL MARCO COMÚN EUROPEO DE REFERENCIA PARA LAS LENGUAS (MCER)

Daniele Zuccalà

Directoras

Dra. Gloria Corpas Pastor Dra. Carla Marello

CORPAS PASTOR GLORIA - 25078844N

Firmado digitalmente por CORPAS PASTOR GLORIA - 25078844N
Fecha: 2023.02.15 09:53:57 +01'00'

Málaga, 2023

DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS PRESENTADA PARA OBTENER EL TÍTULO DE DOCTOR

D./Dña DANIELE ZUCCALÀ
Estudiante del programa de doctorado LINGÜÍSTICA, LITERATURA Y TRADUCCIÓN de la Universidad de Málaga, autor/a de la tesis, presentada para la obtención del título de doctor por la Universidad de Málaga, titulada: INDICADORES FRASEODIDÁCTICOS: UNA PROPUESTA DE SISTEMATIZACIÓN DE LA FRASEOLOGÍA ITALIANA SEGÚN EL MARCO COMÚN EUROPEO DE REFERENCIA PARA LAS LENGUAS (MCER).

Realizada bajo la tutorización de GLORIA CORPAS PASTOR y dirección de GLORIA CORPAS PASTOR, CARLA MARELLO (si tuviera varios directores deberá hacer constar el nombre de todos)

DECLARO QUE:

La tesis presentada es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, conforme al ordenamiento jurídico vigente (Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), modificado por la Ley 2/2019, de 1 de marzo.

Igualmente asumo, ante a la Universidad de Málaga y ante cualquier otra instancia, la responsabilidad que pudiera derivarse en caso de plagio de contenidos en la tesis presentada, conforme al ordenamiento jurídico vigente.

En Málaga, a 05 de FEBRERO de 2023

Fdo.: DANIELE ZUCCA LÀ Doctorando/a	CORPAS PASTOR GLORIA - 25078844N Firmado digitalmente por CORPAS PASTOR GLORIA - 25078844N Fecha: 2023.02.15 09:54:22 +01'00' Fdo.: GLORIA CORPAS PASTOR Tutor/a
CORPAS PASTOR GLORIA - 25078844N Firmado digitalmente por CORPAS PASTOR GLORIA - 25078844N Fecha: 2023.04.28 11:50:57 +02'00' Fdo.: GLORIA CORPAS PASTOR	- CARLA MARELLO

Málaga a 05 de febrero de 2023

Gloria Corpas Pastor, profesora de la Facultad de Filosofía y Letras (Universidad de Málaga),

HACE CONSTAR

Que Daniele Zuccalà, con NIE Y1908634Q es estudiante de doctorado del Programa de Doctorado "Lingüística, Literatura y Traducción", con matrícula activa, y que ha realizado bajo mi dirección, la Tesis Doctoral titulada

"'Indicadores fraseodidácticos: una propuesta de sistematización de la fraseología italiana según el Marco Común Europeo de Referencia para las lenguas (MCER)"

Revisado el presente trabajo estimo que reúne los requisitos establecidos según la normativa vigente. Por lo tanto, **AUTORIZO** la admisión a trámite y defensa pública de esta Tesis Doctoral para optar al grado de Doctor en la Universidad de Málaga.

Y para que así conste, lo firmo en Málaga a 05 de febrero de 2023

CORPAS PASTOR GLORIA - 25078844N

Firmado digitalmente por CORPAS PASTOR GLORIA - 25078844N
Fecha: 2023.02.15 09:54:46 +01'00'

Fdo. Nombre y apellidos del director de la tesis doctoral

Turín a 05 de enero de 2023

Carla Marello, profesora del Dipartimento di Lingue e Letterature straniere e Culture moderne (Università degli Studi di Torino),

HACE CONSTAR

Que Daniele Zuccalà, con NIE Y1908634Q es estudiante de doctorado del Programa de Doctorado "Lingüística, Literatura y Traducción", con matrícula activa, y que ha realizado bajo mi dirección, la Tesis Doctoral titulada

"'Indicadores fraseodidácticos: una propuesta de sistematización de la fraseología italiana según el Marco Común Europeo de Referencia para las lenguas (MCER)"

Revisado el presente trabajo estimo que reúne los requisitos establecidos según la normativa vigente. Por lo tanto, **AUTORIZO** la admisión a trámite y defensa pública de esta Tesis Doctoral para optar al grado de Doctor en la Universidad de Málaga.

Y para que así conste, lo firmo en Málaga a 05 de enero de 2023

Fdo. Nombre y apellidos del director de la tesis doctoral

Carla Marello



Campus de Teatinos s/n. 29071 Málaga
Tel.: 952 13 16 83/1684/1685/1687/3432/3435 - Fax: 952 13 18 23

05 de febrero de 2023

Gloria Corpas Pastor y Carla Marello,

HANCEN CONSTAR QUE

Las publicaciones en coautoría que avalan la tesis doctoral titulada «Indicadores fraseodidácticos: una propuesta de sistematización de la fraseología italiana según el Marco Común Europeo de Referencia para las lenguas (MCER)», y realizada bajo nuestra dirección, no han sido utilizadas en tesis anteriores.

Gloria Corpas Pastor

CORPAS PASTOR GLORIA - 25078844N

Firmado digitalmente por CORPAS PASTOR GLORIA - 25078844N
Fecha: 2023.02.15 09:55:08 +01'00'

Carla Marello



INDICADORES FRASEODIDÁCTICOS:
UNA PROPUESTA DE SISTEMATIZACIÓN DE LA FRASEOLOGÍA ITALIANA
SEGÚN EL MARCO COMÚN EUROPEO DE REFERENCIA
PARA LAS LENGUAS (MCER)

CAPÍTULO I
INTRODUCCIÓN 7
1.1. Introducción 7
1.2. Necesidad y oportunidad de la investigación 9
1.3. Formulación de las hipótesis 11
1.4. Objetivos de la investigación 12
1.4.1. Objetivos generales 12
1.4.2. Objetivos específicos 12
1.5. Estructura de la tesis doctoral 13

CAPÍTULO II
LA FRASEOLOGÍA EN LA LINGÜÍSTICA ESPAÑOLA E ITALIANA 18
2.1. Introducción 18
2.2. Aproximaciones a la fraseología 20
2.2.1. Definiciones terminológicas en los estudios de fraseología 22
2.2.2. Propiedades lingüísticas de las unidades fraseológicas 24
2.2.2.1. Frecuencia 25
2.2.2.2. Institucionalización 25
2.2.2.3. Fijación fraseológica 26
2.2.2.4. Idiomaticidad 26
2.2.2.5. Motivación y transparencia 28
2.2.2.6. Economía lingüística 29
2.2.3. La competencia fraseológica 29
2.3. La fraseología en la lingüística hispánica 32
2.3.1. Principales aproximaciones a una clasificación de las UF 33
2.3.1.1. Julio Casares 34
2.3.1.2. Alberto Zuluaga 41

2.3.1.3. Gloria Corpas Pastor 46
2.3.1.4. Leonor Ruiz Gurillo.....56
2.3.1.5. Mario García-Page Sánchez 62
2.3.1.6. Otras propuestas.....65
2.4. La fraseología en la lingüística italiana.....74
2.4.1. Principales aproximaciones a una clasificación de las UF.....80
2.4.1.1. Simonetta Vietri 81
2.4.1.2. Federica Casadei 83
2.4.1.3. Miriam Voghera 86
2.4.1.4. Raffaele Simone.....91
2.4.1.5. Otras propuestas.....93

CAPÍTULO III
FRASEODIDÁCTICA: ÓRIGENES Y DESARROLLO 105
3.1. Introducción 105
3.2. Orígenes de la fraseodidáctica 109
3.3. Marco Común Europeo de las Lenguas.....111
3.3.1. Sobre el tratamiento de la fraseología en el MCER 115
3.4. Aplicación de la fraseodidáctica: «cuándo».....120
3.5. Dificultades de la fraseodidáctica en las clases de LE 125
3.6. La fraseodidáctica en manuales de ILE y ELE 130
3.7. Análisis de dos métodos didácticos para impartir fraseología en las clases de ILE.....136
3.7.1. Zamora, Pablo; Arianna, Alessandro; Eleonora, Ioppoli; Federica, Simone (2006): Hai voluto la bicicletta…, Perugia, Guerra Edizioni 137
3.7.2. Aprile, Gianluca (2008): Italiano per modo di dire, Firenze, Alma Edizioni.....141
3.8. La fraseodidáctica italiana y española a través de los diccionarios 145
3.8.1. Diccionarios monolingües españoles 146
3.8.2. Diccionarios monolingües italianos.....147
3.8.3. Diccionarios bilingües italiano-español 151

CAPÍTULO IV
INDICADORES FRASEODIDÁCTICOS Y APLICACIONES PRÁCTICAS 156

4.1. Introducción 156
4.2. Agrupaciones habituales de los IF 158
4.2.1. Léxico clave 159
4.2.2. Campos semánticos 164
4.2.3. Unidades didácticas 168
4.2.4. Uso contrastivo 170
4.2.5. Funciones comunicativas 175
4.2.6. Idiomaticidad y equivalencia 179
4.2.7. Motivación 185
4.2.8. Frecuencia 191
4.2.9. Los IF en breve 199
4.3. Principales propuestas metodológicas para enseñar fraseología como LE 204
4.3.1. Contrastividad 206
4.3.2. Léxico 212
4.3.3. Funciones comunicativas 217
4.3.4. Misceláneas 221
4.3.5. Propuestas fraseodidácticas en breve 232

CAPÍTULO V
METODOLOGÍA EMPLEADA 236
5.1. Introducción 236
5.2. Segunda fase: Extracción y anotación fraseológica 239
5.2.1. Elección de un corpus representativo 241
5.2.1.1. Nuovo Espresso, Alma Edizioni 243
5.2.1.2. Il Nuovo Affresco Italiano, Mondadori Education 247
5.2.2. Extracción y anotación de las UF 251
5.2.2.1. Pretest: extracción automática con SynchroTerm y Sketch Engine 258
5.2.2.2. Identificación y extracción manual de la fraseología 261
5.2.2.3. Anotación de la fraseología en el archivo fraseológico 265
5.2.3. Presentación de la fraseología extraída 270
5.2.3.1. Nuovo Espresso 271
5.2.3.2. Il Nuovo Affresco Italiano 274

5.3. Tercera fase: Análisis de la frecuencia en el Focus corpus 278
5.3.1. Reference corpus 281
5.3.2. Focus corpus 287
5.3.3. Propuesta para la creación de la herramienta 289
5.3.3.1. Riesgos y fortalezas de la herramienta 291
5.4. Cuarta fase: Aplicación de los IF al Reference corpus 295

CAPÍTULO VI
CONCLUSIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN 299
6.1. Conclusiones 299
6.2. Futuras líneas de investigación 303

CAPÍTULO VII
REFERENCIAS BIBLIOGRÁFICAS, INDICE DE ANEXOS Y RESÚMEN EN ITALIANO 307
7.1. Referencias bibliográficas 307
7.2. Índice de anexos (CD) 340
7.2.1. Archivo fraseológico_1 340
7.2.2. Archivo fraseológico_2 340
7.3. Resumen en italiano 341
7.3.1. Contenuti e obiettivi della tesi 341
7.3.2. Metodologia seguita 344
7.3.3. Rischi e limiti della metodologia 351
7.3.4. Considerazioni finali 354
7.3.5. Future linee di ricerca 357
7.3.6. Conclusioni 358

CAPÍTULO I

INTRODUCCIÓN

1.1. Introducción

La tesis doctoral «Indicadores fraseodidácticos: Una propuesta de sistematización de la fraseología italiana según el MCER» se centra en la fraseodidáctica del italiano como lengua extranjera, si bien bebe de otras áreas de conocimiento. En particular, el estudio se inscribe en el ámbito de la fraseología y de la fraseodidáctica, centrando su atención en italiano como lengua extranjera. Tomando como referencia las enseñanzas de Wilkins (1977: 111), «sin la gramática se puede transmitir muy poco; sin el vocabulario no se puede transmitir nada[1]» se detecta la importancia de los estudios sobre unidades fraseológicas (UF) en la lingüística aplicada y de corpus.

Por otra parte, la didáctica de las lenguas es un hito fundamental tanto para los hablantes de un idioma extranjero como para los futuros profesionales de traducción, lo que debe reflejarse en el presente estudio. En este sentido, la aproximación a la fraseología de una lengua, en este caso del italiano como lengua extranjera (ILE en adelante), genera desafíos pragmáticos en la didáctica de las UF.

El trabajo de fin de máster defendido por el doctorando durante el curso académico 2015/2016 para el Máster en Traducción para el Mundo

[1] «without grammar very little can be conveyed; without vocabulary nothing can be conveyed». Traducción propia.

Editorial titulado «Estudio comparado de la fraseodidáctica del italiano como lengua extranjera en Málaga (España) y Quito (Ecuador)», demostró que la didáctica del ILE no dedica suficiente atención a dichas unidades fraseológica por diferentes razones que podrían resumirse en una falta de preparación específica del profesorado con respecto a la fraseología por un lado, y, por otra parte, una clara ausencia de indicadores que permitan definir una taxonomía de las mismas que permitiese a los docentes de ILE impartir las unidades fraseologías correctas en los diferentes niveles de referencia del Marco Común Europeo de Referencia para las lenguas: aprendizaje, enseñanza, evaluación (MCER en adelante).

Se pretende, por lo tanto, examinar las unidades fraseológicas que pertenecen a la lengua italiana contemporánea para, en última estancia, establecer unos indicadores fraseodidácticos (IF)[2] verificables que permitan desglosar las diferentes UF y proponer la elaboración de una clasificación de las UF que sirva de apoyo didáctico para el profesorado de ILE y de los profesionales de la traducción.

Para alcanzar dicho propósito, y como se explicará pormenorizadamente en la metodología, se dividirá el estudio en cuatro fases secuenciales exploratorias: revisión bibliográfica de textos de fraseología y fraseodidáctica del italiano y el español; extracción de la fraseología italiana a partir de dos series completas de recientes manuales de ILE; propuesta de un método apto para la comparación de las UF extraídas con un corpus de referencia y, finalmente, la propuesta de aplicación de los IF a al corpus compuesto por las unidades fraseológicas, lo que permitirá proporcionar una propuesta para distribuir las UF en función de los niveles del MCER.

2 Para la definición del término «indicador fraseodidáctico» véase apartado 4.1.

1.2. Necesidad y oportunidad de la investigación

Los avances realizados en fraseología en los últimos sesenta años, el creciente interés registrado a partir de los años noventa del siglo pasado, tanto a nivel nacional como internacional, y la consolidación de revistas, congresos y grupos de investigación en un gran número de lenguas son solo algunos de las muestras que describen el auge que los estudios fraseológicos han tenido en las últimas décadas. Una disciplina joven, aunque con raíces fuertes, que ha cuestionado por primera vez el ideal de construcción libre del discurso: cuando elaboramos un texto somos menos libres de lo que creemos.

Como señalan varios autores (Casares, 1950; Corpas Pastor, 1996; García-Page, 2008), la arquitectura que plasma las lenguas es, en algunos casos, el fruto de una combinación fija de palabras, estructuras prefabricadas que se repiten una y otra vez en el discurso y que han sido asimiladas en nuestro lexicón mental. El cimiento que une el léxico para darle forma y significado es el objeto de estudio de la fraseología y, por tanto, interés principal de la presente tesis doctoral.

Parafraseando a Mel'čuk (1998), los hablantes utilizamos más frases compuestas que palabras sueltas para comunicar y existen ciertas combinaciones previamente formuladas que se repiten en los textos, ya sean orales o escritos. En este sentido, el léxico asume un significado proprio solo dentro de una oración y, por lo tanto, es en ella que la que encuentra su razón de ser. No es lo mismo *tomar leche* que *ser la leche*. La relación que se instaura entre los vocablos y la manera en la que afecta al significado de la oración es el marco del presente estudio doctoral, es decir, la fraseología entendida como subdisciplina de la lexicología (Corpas Pastor, 1996).

La reiteración de estructuras prefabricadas en los textos producidos por hablantes en los distintos idiomas hace imprescindible, tanto para el alumnado de una lengua extranjera como para la formación de futuros

traductores, conocer estos elementos para emplearlos correctamente en sus respectivos ámbitos. Estas estructuras denominadas «unidades fraseológicas» (Zuluaga, 1980; Corpas Pastor, 1996; García-Page, 2008;) no solo representan el *heritage* (patrimonio cultural) de una lengua, sino que, parafraseando a Corpas Pastor (1996), son indicadores de la condición nativa de un hablante. Así, se parte de la definición de las unidades fraseológicas como «unidades léxicas formadas por más de dos palabras gráficas es su límite inferior, cuyo límite superior se sitúa en el nivel de la oración compuesta» (Corpas Pastor, 1996: 20) y que presentan características específicas como la frecuencia de uso, la institucionalización, entendida como fijación de la unidad en la legua, y la idiomaticidad.

Si bien la fraseología española ha sido abordada con profundo pragmatismo, «los estudios de fraseología italiana no han tenido un desarrollo paralelo a los de fraseología española» (Quiroga, 2006: 202). A pesar de que en las últimas décadas se ha despertado el interés de los fraseólogos italianos (Marello, 1989; Vietri, 1995; Casadei, 1996; De Mauro, 1999; Quiroga, 2006), no existe una clasificación aceptada por la mayoría de los estudiosos y, paralelamente, se observa cierta confusión que concierne a la terminología empleada para definir el amplio universo fraseológico.

En la tradición italiana la fraseología no se ha estudiado como disciplina autónoma sino «principalmente como fenómenos morfológicos» (Quiroga, 2006: 202), es decir, como procesos productivos en la formación de palabras. Este déficit que se produce en los estudios fraseológicos de la lengua italiana puede tener repercusiones diversas en las clases de italiano como lengua extranjera y, en particular, en la manera de impartir las UF, lo cual merece un estudio en profundidad. Además, a pesar de que la fraseología es un fenómeno lingüístico común en todas las lenguas y el italiano es un idioma rico en términos fraseológicos (Quiroga, 2006), los estudios que se dedican al análisis de la didáctica de estas

expresiones con este idioma son exiguos tanto desde un punto de práctico como teórico.

En relación con la didáctica de la fraseología italiana, se puede afirmar que su sistematización no ha recibido el interés académico que merece, afectando al estudio de corpus fraseológicos para la formación de estudiantes de ILE y de traductores, como apunta Fernández Prieto (2004). Por todo ello, es imprescindible estudiar la didáctica de la fraseología del italiano o, utilizando la terminología empleada por Larreta Zulategui (2001) y González Rey (2012), de la fraseodidáctica. Del mismo modo, la escasez de literatura académica sobre fraseodidáctica del italiano hace que la presente tesis doctoral sea un estudio necesario y, al mismo tiempo, tenga un carácter innovador.

1.3. Formulación de las hipótesis

La presente tesis doctoral parte las siguientes hipótesis:

H1. La mayoría de las propuestas para aplicar los indicadores fraseodidácticos a la fraseología italiana sufren de cierto grado de subjetividad y los corpus fraseológicos que emplean no son compilados de manera empírica.

H2. La frecuencia es uno de los indicadores fraseodidácticos más eficaz a la hora de asignar la fraseología a cada nivel del MCER.

H3. La extracción manual de la fraseología proporciona más resultados fiables que la automática.

En consecuencia, se pretende verificar o refutar las hipótesis definidas anteriormente a través de una investigación que corrobore si realmente la fraseodidáctica puede ser abordada de manera empírica en las aulas de italiano como lengua extranjera.

1.4. Objetivos de la investigación

A continuación, se definen los dos objetivos generales y los cinco específicos sobre los que la presente tesis doctoral pone sus fundamentas.

1.4.1. Objetivos generales

O1. Estudiar la fraseología italiana aplicada a la didáctica del ILE.

O2. Identificar las principales propuestas académicas sobre los indicadores fraseodidácticos en ILE y ELE.

O3. Extraer la fraseología objeto de estudio a partir dos series completas de manuales de ILE y estudiar su composición.

1.4.2. Objetivos específicos

O4. Identificar los Indicadores fraseodidácticos (IF) más significativos para el italiano como lengua extranjera (ILE).

O5. Proponer un modelo para compilar un corpus fraseológico adaptado al MCER que se base en los indicadores más significativos para el ILE.

O6. Elaborar y anotar la fraseología en un archivo electrónico *ad hoc* siguiendo estándares internacionales.

O7. Diseñar una propuesta de una herramienta que permita extraer la fraseología de un corpus ad hoc con uno de la lengua general.

Para concretar dichas metas, el presente estudio se divide en seis capítulos, así como se han definido en el apartado 1.2. Diseño de la tesis doctoral.

1.5. Estructura de la tesis doctoral

El diseño de la presente tesis doctoral se ha desarrollado con el fin de responder a interrogantes sobre la fiabilidad de las propuestas prácticas para llevar la fraseología al aula de italiano como lengua extranjera. En este apartado se presentan las estrategias básicas que se han empleado para obtener información verídica y empírica sobre el objeto de estudio y se describe la estructura general de la presente investigación.

La presente tesis doctoral se estructura en siete capítulos.

En el capítulo I se detallan las necesidades y oportunidades del estudio, se describe de manera general el contenido de los capítulos que componen la tesis y, finalmente, se definen las hipótesis y los objetivos que guían la investigación.

En segundo lugar, y con el objetivo de profundizar el objeto de estudio antes de empezar la investigación, se detallará un profundo análisis documental que por un lado permitirá enmarcar la investigación en el espacio y el tiempo y, por otro lado, analizar críticamente la literatura previa a dicho trabajo (Capítulo II). En este sentido, se analizará pormenorizadamente la literatura que ha dado lugar a las taxonomías de la fraseología en dos idiomas, el español y el italiano. El hecho de que un estudio sobre la fraseología italiana analice también fuente primarias y segundarias sobre fraseología española no es baladí, sino que, como se verá a lo largo de la presente investigación, es justificado por la significativa influencia que dichas aportaciones has tenido sobre los estudios italianos a lo largo de las últimas décadas.

En el siguiente capítulo (capítulo III), el estudio se centra en el estado de la cuestión sobre el concepto de fraseodidáctica y se enmarcan los orígenes de dicha área de conocimiento y se define el concepto. Asimismo, se analiza el MCER desde una perspectiva meramente fraseológica con el objetivo de comprender qué tratamiento se reserva

a la fraseología y, por ende, qué líneas guías ofrece al profesorado que quiere trabajar la fraseología en los diferentes niveles del MCER.

Posteriormente se presentan numerosas propuestas científicas para guiar al profesorado en decidir en qué momento del ciclo formativo del alumnado se debería de incluir la fraseología en las clases de lengua extranjera. En este sentido, se analizan las principales corrientes que defienden o descartan la introducción de la fraseología a partir de los niveles más básicos de MCER (A1, A2). Asimismo, se presentan las principales dificultades que el profesorado encuentra a la hora de impartir fraseología en las aulas de LE y, para ello, además de analizar fuentes primarias y secundarias, se recurre al estudio que el doctorando presentó como trabajo fin de máster en el que profundizó en dichas dificultades a través de un estudio de caso.

A continuación, se analiza el tratamiento de la fraseodidáctica en los manuales de italiano como lengua extranjera (ILE) y de español como lengua extranjera (ELE) para posteriormente analizar en profundidad dos casos de estudios para impartir la fraseología en las clases de ILE: uno en el que la fraseología es tratada de manera pragmática y científica, y otro que no representaría un modelo incompleto y subjetivo para trabajar la fraseología en el aula. Finalmente, el último apartado de dicho capítulo es dedicado al análisis de la fraseodidáctica italiana y española en los diccionarios: monolingües españoles, monolingües italianos y bilingües italiano-español.

El capítulo IV de la presente investigación se centra en los que se han definido con el término de «Indicadores fraseodidácticos» (IF), es decir, un conjunto de herramientas diagnósticas, tanto cualitativas como cuantitativas, que permiten didactizar las unidades fraseológicas y cuya aplicación permite didactizar la fraseología en función del MCER desde un prisma didáctico. En este sentido, en primer lugar, se identificarán los principales IF a través de un atento análisis de las propuestas previas al

presente estudio y, en segundo lugar, se analizarán las principales propuestas metodológicas para aplicar dichos indicadores en las clases de lengua extranjera.

En el siguiente capítulo V se define el diseño metodológico que se ha empleado en el presente trabajo doctoral. En este sentido, se presentan los orígenes del modelo secuencial exploratorio propuesto por el profesor Creswell (2003, 2007, 2009, 2014) y se define una metodología de tipo cualitativo y cuantitativo que trata de conseguir la mayor información sobre el objeto de estudio mediante distintas técnicas aplicadas de manera secuencial. Siguiendo dicho modelo, la metodología de esta investigación ha sido dividida en cuatro fases.

Tras la profunda revisión documental (fase 1) previamente descrita, se procederá con la extracción de la fraseología en dos series completas de manuales de ILE (fase 2); para ello se describirán en primer lugar las características del corpus representativo elegido: dos series completas (del nivel A1 al C2) de dos métodos de referencia para la didáctica del italiano como lengua extranjera (ILE). Con el objetivo de averiguar si la extracción automática de la fraseología a partir de dicho corpus pudiese ser cuantitativa y cualitativamente válida, se presentará un pretest realizado a través de tres herramientas de extracción automática (SynchroTerm, Sketch Engine y el Term Extraccion Module de IATE, la base terminológica de la Unión Europea). Puesto que los resultados fueron decepcionantes, se optó por realizar una extracción manual de fraseología.

Asimismo, se definirá el método empleado basado en estándares internacionales para la anotación de la UF en un archivo fraseológico[3]

[3] En el presente estudio, el término archivo fraseológico se usa con significado asimilable a colección/listado electrónico de UF recopiladas y organizadas a partir de criterios comunes. El termino archivo fraseológico se toma prestado a partir de fichero terminológico: «Conjunto de fichas terminológicas que están relacionadas lógicamente en una base de datos mediante el uso de un mismo modelo de presentación, la aplicación de las mismas reglas de consignación y la utilización de un único modelo de consulta.» (Pavel y Nolet, 2001: 127). En este sentido, con archivo fraseológico se entiende un

elaborado *ad hoc* en formato electrónico para el presente trabajo. La presentación desde un prisma compositivo de fraseología extraída concluirá la última parte de la segunda fase de la metodología.

A continuación (fase 3), se presentará un modelo secuencial para identificar la frecuencia de las unidades fraseológicas previamente extraída (*Refence corpus*) en un corpus general de la lengua (*Focus Corpus*). Para ello se presentará la propuesta para la creación de una herramienta *ad hoc* realizada en colaboración con el *IATE Support & Development Team*, entidad interinstitucional con la que el doctorando estuvo trabajando en Centro de Traducción de la Unión Europea durante ocho meses. Un análisis estratégico de los riesgos y las fortalezas de dicha herramienta será presentado en el siguiente apartado con la idea de limitar las amenazas y fortalecer las oportunidades.

Dicha herramienta, basada en los principios del procesamiento de lenguaje natural (PLN) y de extracción terminológica, permitirá acceder a la última fase del diseño metodológico (fase 4), es decir, la aplicación de los indicadores fraseodidácticos al corpus compuesto por las unidades fraseológicas. En este sentido, se propone aplicar los resultados de la fase anterior a los distintos niveles del MCER y definir el nivel jerárquico que ocupa cada unidad extraída dentro de un nivel de referencia del MCER.

Finalmente, el último capítulo (capítulo VI) presentará las conclusiones y las consideraciones finales que pretenden ser el principio de las futuras líneas de investigación. Las referencias bibliográficas cerraran la tesis doctoral que se presenta en las próximas páginas.

archivo en-linea (formato .xlsx) que presenta un listado de UF ordenadas mediante el mismo criterio y modelo.

CAPÍTULO II

LA FRASEOLOGÍA EN LA LINGÜÍSTICA ESPAÑOLA E ITALIANA

2.1. Introducción

La fraseología se podría caracterizar por ser un área líquida de la lingüística, para emplear el concepto de Zygmunt Bauman (2000). El sociólogo y humanista polaco empleó la locución «sociedad líquida» para definir el actual momento histórico en el que las realidades sólidas de nuestros antepasados se han desvanecido, dejando espacio a un mundo más precario, provisional y, a menudo, difícil de encasillar. Algo similar ocurre con la fraseología: en comparación con el léxico simple, las propiedades intrínsecas de la fraseología hacen que huya de etiquetas y cuadros preestablecidos, se escabulle, esquiva los encasillamientos. Por esta misma razón, y como se verá a lo largo del capítulo, las unidades fraseológicas (UF) han despertado el interés de lexicógrafos y terminológicos que, atraídos por sus propiedades, han intentado con más o menos éxito encasillar esta área de conocimiento en clasificaciones de diferente naturaleza.

En el presente capítulo se abordarán dos temas que han movido los estudios fraseológicos hasta el día de hoy: en primer lugar, las características intrínsecas de las unidades fraseológicas y, en segundo lugar, las principales taxonomías que se han elaborado en el idioma español e italiano. Así, se procederá en un primer apartado al análisis de las características y propiedades fraseológicas de ambos idiomas. Dichas propiedades contraponen las unidades léxicas simples a las complejas puesto que, si por un lado disponen de características comunes, por el otro poseen rasgos independientes.

Con respecto al segundo punto, a lo largo del capítulo se analizarán las principales aportaciones que han abordado la difícil tarea de organizar todo el universo fraseológico en una única clasificación. Para ello, se presentarán pormenorizadamente las aproximaciones españolas e italianas y se observará que, si bien las primeras han tenido un gran desarrollo en las últimas décadas, las italianas han sido elaboradas de manera diferente dependiendo del momento: en una primera fase desde una perspectiva meramente lexicográfica y, en un segundo momento, se han nutrido de los estudios españoles, aplicándolos como instrumento de comparación.

Finalmente, la fraseología de ambos idiomas está profundamente relacionada a la terminología que se emplea para describir el universo fraseológico. De hecho, «parlare di fraseologia significa definire la fraseologia stessa e cercare di individuare la terminologia che la caratterizza»[4] (De Giovanni, 2015). Por ello, se presentan las diferentes propuestas terminológicas, describiendo tanto las más frecuentes, como las que han recibido mayor aceptación a lo largo de la breve historia de la fraseología, evidenciando la vasta nomenclatura utilizada en español y la cierta discrecionalidad que sufre, mientras que, en italiano, se observa literalmente un «caos terminológico» (Mellado Blanco, 2010: 358) que merece ser estudiado para llegar a un compromiso entre las diferentes propuestas.

Con el objetivo de definir el marco teórico de la presente tesis doctoral, a lo largo de este capítulo se profundizará en el concepto de fraseología entendida desde una perspectiva amplia, cuyo objeto de estudio son los segmentos oracionales compuestos por más de una palabra, pero cuya dependencia semántica es tan fuerte y estrecha que se presenta a la

[4] «Hablar de fraseología significa definir la propia fraseología e intentar identificar la terminología que la caracteriza». Traducción propia.

«coscienza dei parlanti come se fossero una parola sola»[5] (Marello, 2008 [1996]: 2).

2.2. Aproximaciones a la fraseología

La fraseología posee una estrecha relación con la terminología y, al mismo tiempo, con la lexicografía, puesto que su aplicación no se limita exclusivamente a las lenguas especializadas, sino que incluye también la lengua común, es decir, la que se utiliza en la vida cotidiana. El fraseólogo es el especialista de esta disciplina, de igual modo que el terminólogo es el especialista en terminología. Es por lo tanto indispensable, en primer lugar, disponer de sólidas nociones en terminología y definir los conceptos que se emplean en la presente tesis doctoral.

Silvia Padel y Diane Nolet, en su ya clásico *Manual de terminología*, definen la terminología como la «disciplina lingüística dedicada al estudio científico de los conceptos y términos utilizados en las lenguas de especialidad» (xvii, 2001). En otras palabras, es la ciencia que forma parte de la lingüística aplicada y que comprende el trabajo en lexicografía especializada, traducción, redacción y enseñanza de idiomas. La *conditio sine qua non* para abordar estudios de terminología y, por lo tanto, de fraseología, es conocer en detalle las normas internacionales que priman en terminología y, en modo particular, el ISO 704:2000, *Terminology work. Principle and methods*[6]. El ISO 704 es la norma que estructura y estandariza los elementos esenciales para el trabajo cuantitativo en terminología (ISO 704, vi: 2000). En este sentido, El ISO 704:

> establishes and harmonizes the basic principle and methods for preparing and compiling terminologies both inside and outside the framework of standardization [...] is applicable to terminology work

[5] «Conciencia de los hablantes como si fueran una palabra sola». Traducción propia.

[6] Segunda edición, 15/11/2000.

> in scientific, technological, industrial, administrative, and other fields of knowledge.[7] (ISO 704, 1:2000)

Cabe por lo tanto definir la nomenclatura terminológica básica que se repetirá a lo largo del trabajo doctoral: término, concepto y objeto. Para ello, hay que acudir a otro estándar internacional: el ISO 1087-1:2000 *Terminology work -- Vocabulary -- Part 1: Theory and application*[8].

En él, se define un *término* como «verbal designation of a general concept in a special field» (6: 2000), es decir, la denominación verbal de un concepto genérico en un dominio específico de la lengua. Un término puede incluir letras, símbolos, numerales, símbolos matemáticos, signos ortográficos y sintácticos como paréntesis, puntos, etc. Los términos que en una misma lengua representan el mismo concepto son sinónimos; los que representan el mismo concepto en dos lenguas diferentes son equivalentes. Asimismo, pueden presentar variantes morfológicas y ser simples o complejos, si presentan dos o más palabras. En este sentido, la presente tesis doctoral emplea el concepto de unidades fraseológicas como sinónimo de términos complejos.

Por otra parte, cabe definir la diferencia entre *concepto* y *objeto*[9]. Los conceptos son definidos por la International Standard Organization como «units of knowledge created by a unique combination of characteristics[10]» (ISO 1087-1, 2:2000) y se consideran representaciones mentales de un objeto en un contexto definido. Son, por ende, la abstracción basada en el conjunto de los objetos que toman parte a su

[7] «Establece y armoniza el principio básico y los métodos de preparación y compilación de terminología tanto dentro como fuera del marco de la estandarización [...] es aplicable a la labor terminológica en los ámbitos científico, tecnológico, industrial, administrativo y otros campos del conocimiento». Traducción propia.

[8] El ISO 1087-1:2000 define y estandariza el vocabulario básico para la teoría y aplicación del trabajo terminológico.

[9] Para más información sobre los tipos de conceptos, las características y las relaciones que se instauran entre ellos o entre objetos y conceptos, véase el ISO 704 (2-15: 2000) e ISO 1087-1 (2 y ss.: 2000).

[10] «unidades de conocimiento creadas a través de una combinación única de características». Traducción propia.

formación y no puede existir como unidad aislada, sino solo en relación con otras unidades, es decir, «son representaciones mentales que ayudan a estructurar los objetos en el mundo real» (Pavel y Nolet, 23: 2001).

Por ello, no se debe confundir los conceptos con los objetos, que son la representación gráfica y concreta de los primeros, puesto que «are assumed to exist and attention shall be focused on how one deals with objects for purpose of communication [...] an object is defined as anything perceived or conceived11» (ISO 1087-1, 2: 2000). Por último, los objetos pueden ser objetos concretos como, por ejemplo, un diamante o un río; inmateriales o abstractos, como la gravedad o un plan financiero; o puramente imaginarios, como la piedra filosofal o un unicornio. Estos tres elementos son la base de los estudios de fraseología teniendo estas estrechas relaciones con la terminología y la lexicografía.

2.2.1. Definiciones terminológicas en los estudios de fraseología

A pesar de la breve historia de la fraseología (Alessandro, 2011), en los estudios académicos que abordan esta disciplina existe una gran disparidad en cuanto a la terminología específica empleada para denominar el objeto de estudio del presente trabajo. Tanto en español como en italiano, se han propuesto diferentes términos por lo que urge, en primer lugar, realizar un esfuerzo por recopilar esta terminología y, en segundo, justificar cuáles de estos términos se han empleado en el presente estudio, para identificar, en un tercer momento, las razones que han llevado a tomar dicha decisión. A continuación, se presentan las principales denominaciones para los dos idiomas objeto de esta tesis doctoral.

Con respecto a las investigaciones sobre fraseología española, se han empleado numerosas designaciones para expresar la unidad básica o

[11] «se supone que existen y la atención se centra en cómo se trata a los objetos con fines de comunicación [...] un objeto se define como cualquier cosa percibida o concebida». Traducción propia

término complejo, objeto de estudio de la fraseología. Entre ellas se menciona *expresión pluriverbal* (Casares, 1950), *expresión fija* (Zuluaga, 1980), *unidades pluriverbales* (Casares, 1950; Corpas, 1996; García-Page, 2008), *unidades polixemáticas* (Cabré, 1996), *modismo* (Casares, 1950); *fraseologismo* (Mellado 2004), *frasema*, (Martí Sánchez, 2005), etc. El MCER en su traducción realizada por el Instituto Cervantes (2002), no solo no resuelve este conflicto terminológico, sino que lo agrava, empleando terminología diferente para definir el mismo objeto de estudio.

Sin embargo, el término más ampliamente utilizado en los ensayos en español sobre fraseología es *unidad fraseológica*, utilizado por numerosos autores[12] como hiperónimo de todas las clases fraseológicas que se examinan en esta investigación. El primer autor en emplear dicha designación es Alberto Zuluaga (1980), pero es con Corpas Pastor que dicha designación se consagra, poniendo de acuerdo gran parte de la comunidad científica en su ya clásico *Manual de fraseología española* (1996).

En la presente tesis doctoral se ha decidido continuar con la línea adoptada por la mayoría de los trabajos académicos en español sobre fraseología, optando por emplear *unidad fraseológica* y su sigla UF y las identifica como: «unidades léxicas formadas por más de dos palabras gráficas en su límite inferior, cuyo límite superior se sitúa en el nivel de la oración compuesta» (Corpas Pastor: 1996: 20) y que presentan una serie de características específicas que se detallarán en los siguientes apartados.

Si en español se ha llegado a cierto consenso académico, en italiano estamos aún lejos de ello. Frente a la falta de terminología aceptada por la mayoría de la academia (Marello, 1996), la terminología empleada es inestable y variada. Las principales propuestas han sido: *espressione*

[12] Para un listado exhaustivo véase Corpas Pastor (1996).

multiparola (Berruto, 1976); *unità lessicale superiore* (Dardano, 1978); *unità* o *espressione polirematica* (De Mauro en la introducción del *GRADIT* 1999-2007; Voghera, 2004), *lessema polirematico* (De Mauro, 1993; Voghera, 1994), *lessema complesso* (De Mauro y Voghera, 1996), *parola complessa* (Simone 2008), *parola sintagmatica* (Masini, 2007), *parole giustapposte* o *sintagmi lessicalizzati* (Marello, 1996), *frase idiomatica* (Elia, 1981; Vietri, 1990), *espressione fissa* (Goody, 1997), el más genérico *frase fatta* (Pittano, 1992); etc.

En este contexto, la denominación recientemente más empleada por los italianistas que se ocupan de fraseología italiana es *espressione idiomatica* (Vietri, 1985, 1990; Casadei, 1996). No obstante, como se verá a continuación, existen numerosos hispanistas que estudian la fraseología italiana y, entre este grupo de académicos, parece más frecuentemente el calco lingüístico del español *unità fraseologica* (UF) (Alessandro, 2006; Núñez Román, 2015). En el presente trabajo doctoral se prioriza esta última designación también para estudios en lengua italiana puesto que, como se verá a lo largo del presente capítulo (2.2.1.3. Gloria Corpas Pastor), la clasificación propuesta por Corpas Pastor (1996) es la que actualmente se ha llevado con más éxito a la fraseología italiana.

2.2.2. *Propiedades lingüísticas de las unidades fraseológicas*

Para definir el objeto de la fraseología, es decir, las unidades fraseológicas, se analizarán brevemente sus principales características. Cabe destacar que el presente apartado no desea presentar un análisis exhaustivo de las características lingüísticas de las UF[13], sino apuntar las más relevantes para el enfoque fraseodidáctico que vertebra el presente estudio.

[13] Las propiedades de las UF han sido ampliamente debatidas y estudiadas ya desde los orígenes de la fraseologìa. Para el análisis de una muestra más detallada véase Casares (1950), Zuluaga (1980), Corpas Pastor (1996), David (2000), Ruiz Gurillo (2001a) y Olza Moreno (2009), entre otros.

2.2.2.1. Frecuencia

Para comprender el concepto de frecuencia es necesario realizar una primera distinción entre *frecuencia de coaparición* y *frecuencia de uso*. La frecuencia de coaparición es «la que presentan aquellas UF cuyos elementos constituyentes aparecen combinados con una frecuencia de aparición conjunta superior a la que cabría esperar según la frecuencia de aparición individual de cada palabra en la lengua» (Corpas Pastor, 1996: 20-21). En cambio, la frecuencia de uso describe cuando determinada combinación es aplicada en el discurso. La frecuencia de uso asume un rol primario en la metodología del presente trabajo, puesto que será uno de los indicadores para establecer qué UF proporcionar a los discentes de italiano como LE. Cabe apuntar que el presente estudio identifica en la frecuencia una de las características denotativas de la fraseología puesto que el uso repetido, así como la coaparición repetida de sus elementos composicionales, dan vida a otros fenómenos que se analizan en esta sección.

2.2.2.2. Institucionalización[14]

Puesto que es imposible llegar a manejar la totalidad del léxico de una lengua (Villarini, 2021), cuanto más alta sea la frecuencia de uso de una determinada UF, es decir, más se utilice en la lengua escrita o hablada, mayores posibilidades se dan de emplearla dentro de una comunidad lingüística. En este sentido, García-Page define la institucionalización como

> El proceso por el cual una comunidad lingüística adopta una expresión fija, la sanciona como algo propio, como moneda de cambio en la comunicación cotidiana, como componente de su acervo lingüístico-cultural, de su código idiomático, como

[14] García-Page (2008) aunque el concepto es definido anteriormente por Corpas Pastor (1996).

> cualquier otro signo convencional, y pasa a formar parte del vocabulario. (García-Page, 2008: 29)

Este no es un proceso inmediato, ya que está estrechamente ligado al uso, la repetición y la frecuencia. En este sentido, la repetición conduce a la cristalización de una expresión y, como consecuencia, a convertirse en una unidad que estará a disposición de los hablantes para su uso. Los hablantes repiten combinaciones ya utilizadas por otros que funcionan como unidades léxicas y, como tal, se almacenan en el lexicón mental.

2.2.2.3. Fijación fraseológica

La fijación fraseológica es el «carácter supremo de la fraseología» (García Page, 2016: 39). Casi la totalidad de los autores han tratado la fijación como uno de los rasgos distintivos de las UF. Para proporcionar una definición de la fijación fraseológica, se emplea aquí la definición de Mellado, quien la entiende como «el resultado directo de su uso repetido dentro de una determinada comunidad lingüística» (2004: 53).

Los procesos de fijación, resumidos por Zuluaga (1980), representan uno de los rasgos principales que distinguen las combinaciones libres de las fijas. En este sentido, la consolidación en el uso hace que las expresiones no idiomáticas se conviertan en idiomáticas y que se establezca una asociación directa entre la UF y la interpretación semántica que le atribuye una comunidad de hablantes. Desde este punto de vista, el uso de una UF implica un gran conocimiento del contexto de uso y de las reglas sociales que la rigen. Por todo ello, los aspectos pragmáticos son rasgos fundamentales para el presente trabajo doctoral y, en general, para los estudios en didáctica de la fraseología.

2.2.2.4. Idiomaticidad

Desde un punto de vista de la gramática de construcciones (CxG del inglés *Construction Grammar*) las construcciones son consideradas

simbólicas en cuanto comprenden el binomio de su forma denotativa particular con su forma connotativa (significado) y/o función específica. En otros términos,

> Las palabras, así como las UF, tienen una forma que se asocia convencionalmente a un polo de significado que es, por defecto, no compositivo (no puede derivarse o computarse a partir del significado unitario de sus componentes). (Corpas Pastor, 2022: 2)

La idiomaticidad es la característica por la cual «el significado global de dicha unidad fraseológica no es deducible del significado aislado de cada uno de sus elementos constitutivos» (Corpas Pastor, 1996: 26). Una expresión idiomática debe ser interpretada como un segmento único e inseparable; en este sentido, el presente estudio considera la lengua como un continuum idiomático cuyos elementos bases son las construcciones idiomáticas (Corpas, 2021). Dicha característica está estrechamente ligada a la cultura de la comunidad de habla que la emplea (Olza, 2009), es decir, a la manera propia de concebir la realidad en la que vive dicho grupo social y, en esta realidad, asume un sentido unitario para una comunidad de hablantes.

Se distinguen así las construcciones no idiomáticas de las idiomáticas. En las primeras, los componentes y la relación que se establece entre ellos conservan su identidad semántica, mientras que en las idiomáticas pierden dicha identidad. No obstante, cabe destacar que no todas las UF son idiomáticas, ya que se trata de «una característica potencial, no esencial de este tipo de unidades» (Corpas Pastor, 1996: 27). Por ello, se han establecido una serie de grados de idiomaticidad[15] que van de las locuciones con un grado alto de idiomaticidad, a aquellas caracterizadas exclusivamente por su fijación, pasando por las

[15] Para profundizar en los grados de idiomaticidad véase: Zuluaga (1980: 129-134) Corpas Pastor (1996: 25-27), Ruiz Gurillo, 1997: 96-99), Pamies Bertrán (2014: 33-46), Sarancho Arnáiz (2015a: 104-107), entre otros.

combinaciones semifijas (Ruiz Gurillo, 1997), que se definirán en los apartados sucesivos.

2.2.2.5. Motivación y transparencia

La motivación y la transparencia están estrechamente ligadas al concepto de idiomaticidad. La motivación se refiere a la existencia de una relación diacrónica entre el significado literal y el figurado de una UF (Pamies, 2014). En este sentido, toda UF posee su motivación en la diacronía aunque esta relación puede haberse perdido a lo largo de la evolución de una lengua y su origen puede no ser reconocido por los hablantes nativos, o ideales, de un idioma. Es un proceso diacrónico y, por ello, de difícil acceso (Heine, 1997: 10-11) y cuando dicha relación se pierde, la unidad fraseológica se hace arbitraria.

La motivación no se debe confundir con la idiomaticidad, ya que «la motivación afecta a la existencia, conocida o no, de una conexión entre la imagen literal y el sentido figurado, ambos tomados en su totalidad» mientras que la segunda «resulta una ecuación entre el todo y las partes» (Pamies Bertrán, 2014: 36). Con todo, una determinada UF puede ser transparente, es decir, su significado se puede deducir de manera más o menos inmediata, u opaca, si este no se puede deducir siendo oscura para los hablantes nativos de un idioma.

La transparencia de una UF no es dada exclusivamente por su motivación diacrónica, sino que a veces puede ser deducida por los hablantes puesto que su significado denotativo es literal en lugar de ser figurativo, es decir, no idiomático. El ISO 704: 2000[16] afirma que un término es considerado transparente cuando «the concept it designates can be inferred, at least partially, without a definition[17]» (2000: 25), es decir, su

[16] Segunda edición 15/11/2000.

[17] «el concepto que designa puede ser deducido, al menos parcialmente, sin una definición». Traducción propia.

significado es comprensible directamente es su morfología. Si un término, simple o complejo, no es transparente, será considerado opaco.

2.2.2.6. Economía lingüística

Las unidades fraseológicas tienen el gran potencial de representar en pocos términos un concepto completo. En este sentido, las UF son concisas y respetan el principio lingüístico de la economía, el principio que gobierna la formación de nuevos términos (ISO, 2000: 25). Frente a una UF demasiado larga los hablantes recurren a estrategias de economía lingüística como la elipsis: eliminan parte de la unidad. Esta situación ocurre mucho con las paremias, donde los hablantes usan, y a menudo recuerdan, solo la primera parte de la misma. Sin embargo, cabe destacar que esto no siempre es así. Hay UF pluriverbales muy extensas que se emplean en su forma completa a pesar de que sean completamente opacas. Asimismo, la elipsis no es privativa de las unidades largas y no es la única manera de manipular o acortar las unidades para crear efectos expresivos[18].

2.2.3. La competencia fraseológica

Desde que la fraseología sentó sus bases como ciencia autónoma e independiente de la lexicografía, muchos estudiosos han ampliado el concepto de competencia lingüística elaborado por Noam Chomsky (1965, 1980) y llevándolo al terreno de la fraseología: competencia fraseológica (CF)[19]. Chomsky, como Saussure, considera que la lengua no es una mera enumeración de vocablos, sino que posee una estructura propia. Sin embargo, a la distinción de Saussure entre *langue*/*parole*, Chomsky propone el binomio *competencia*/*actuación*, donde la primera depende de la capacidad de elaborar una oración, es decir, el conocimiento de la lengua, y la segunda viene a ser la realización

[18] Véase Corpas Pastor (1996: 29), así como los conceptos de manipulación creativas propuestos por la misma autora (1996: 233 y ss.).

[19] En algunos trabajos se utiliza competencia idiomática (CI).

práctica de esta competencia o, en otras palabras, la producción efectiva del hablante-oyente en una situación real.

El concepto de competencia lingüística, en un primer momento entendido exclusivamente desde el punto de vista puramente gramatical, hace referencia al concepto de hablante-oyente ideal. Con ello, Chomsky (citado en Iragui, 2004) se refiere a un:

> Hablante-oyente ideal de una comunidad de habla completamente homogénea que conoce su lengua perfectamente y al que no le afectan condiciones irrelevantes a nivel gramatical como las limitaciones de memoria, las distracciones, los cambios de atención y de interés y los errores al aplicar su conocimiento de la lengua a la actuación real. (Chomsky, 1965: 3)

Este concepto de hablante ideal, nativo, es difícil de aceptar, puesto que supone una abstracción e idealización que no posee una relación directa con la capacidad para utilizar una o más lenguas por parte de hablantes monolingües y plurilingües. En este sentido, dicha concepción es reduccionista si se aplica a la adquisición y enseñanza de lenguas. Por otra parte, el concepto de competencia lingüística es igualmente limitado, al no considerar las otras competencias que un hablante-oyente dispone, como la pragmática, la comunicativa, etc.

Por esta razón, unos años más tarde, Hymes introduce el concepto de competencia comunicativa, es decir, «las reglas de uso sin las cuales las reglas gramaticales serían inútiles» (Hymes, 1972: 278). La competencia se amplía, el uso convive con la lengua y, por supuesto, con otros factores extralingüístico. Cabría suponer que la competencia de un hablante no puede ser restringida a un solo tipo de competencia, sino a una serie de competencias específicas o, en otras palabras, a una multicompetencia.

Es en este contexto donde toma partido el concepto de competencia fraseológica (CF).

Existen numerosos estudios sobre la CF de un hablante nativo o extranjero[20]. Definir el concepto de CF es complejo ya que se ha producido un reducido análisis de esta competencia, si se compara con las demás competencias. Por otra parte, cabe destacar que ha recibido especial atención desde la perspectiva de los especialistas de la didáctica de las lenguas (Martí Sánchez, 2016). De acuerdo, con Hymes (1984), la competencia fraseológica se sitúa bajo la competencia comunicativa.

Por el contrario, el MCER no asume un papel independiente de la CF, sino que se encuentra diseminada en las demás competencias. El Marco Común Europeo de Referencia para las lenguas reconoce las siguientes competencias: lingüísticas (léxica, gramatical y semántica, fonológica, ortoépica y ortográfica), sociolingüística y pragmáticas (discursiva, organizativa y funcional).

Toda esta complejidad en su aproximación complica la definición de la competencia fraseológica. Si se intenta proporcionar una definición inspirada en las primeras teorías *chomskianas* de hablante-oyente ideal, nativo, se hace aún más complicado. Un hablante nativo desconoce cómo descifrar todas las unidades fraseológicas de su lengua, ya sea por su naturaleza, porque han caído en desuso, porque se utilizan en un determinado entorno social, geográfico, cultural o porque pertenecen a un lenguaje específico, y ello no condiciona que continúe siendo el hablante ideal de un idioma.

[20] Como pequeña muestra véase: Claire Nicolas (2008) y Florence Detry (2010) con el proyecto FRASEOTEXT, Julia Sevilla Muñoz y Marina García Yelo es su comunicación «Las paremias en la competencia comunicativa del francés actual con vistas a la enseñanza de lenguas», presentado en el XVII Congreso de la APFUE en 2008, o Natalia Vyshnya (2006) en su estudio «Adquisición de las competencias lingüísticas en las clases del idioma español (L2) a través de las paremias clásicas de El Quijote».

Sin embargo, la CF fraseológica es real, al igual que también es real un conocimiento fraseológico que permite reconocer una UF como tal frente a una asociación libre de palabras. Por ello, Martí Sánchez prefiere hablar de conocimiento fraseológico como de «instrucciones de uso e interpretación como manifestación de un conocimiento fraseológico» (2016: 62-65), considerándolo como equivalente a la competencia fraseológica. Dicho conocimiento es intrínseco, por lo tanto, a una determinada comunidad de habla y la persona que no pertenece a la misma limita significativamente su competencia/actuación, para volver a la concepción de Chomsky.

En la presente tesis doctoral no se pretende proporcionar una definición unívoca de la CF, sino poner en evidencia que se trata de una cuestión que merece un mayor esfuerzo desde la fraseología y la fraseodidáctica, al ser «el desarrollo de la competencia fraseológica» (Alessandro 2011: 211) uno de los fines de esta última. En este sentido, el estudio prolongado de una lengua puede llevar un hablante a conseguir un grado notable de competencia fraseológica, parecido al de una determinada comunidad de habla, pero todavía es imposible afirmar si esta será tan profunda y efectiva como la de un hablante nativo.

2.3. La fraseología en la lingüística hispánica

En el presente apartado se presentarán de manera diacrónica las principales clasificaciones del universo fraseológico elaboradas por hispanistas, así como los debates conceptuales y terminológicos que se han cuestionado desde que dicha disciplina se comenzó a abordar en España. Asimismo, se examinará la amplia terminología empleada por los fraseólogos para definir dichas unidades que, en algunos casos, son coincidentes y, en otros, solo poseen algunos matices en común. Con todo, y como se verá a lo largo del apartado, se puede afirmar que la clasificación que actualmente recibe el mayor consenso por la

comunidad científica es la de Corpas Pastor (1996), y es la que se utiliza como referencia en la presente investigación.

2.3.1. Principales aproximaciones a una clasificación de las UF

Desde hace siglos se registran unidades fraseológicas en el ámbito hispánico mediante obras lexicográficas, gramáticas y otros tipos de textos, principalmente literarios. Los diccionarios han tratado de describirlas, puesto que, como se ha visto las UF son unidades formalmente estables y presentan una cadena de vocablos que permanece prácticamente invariable. Por todo ello, las unidades fraseológicas en general y los proverbios, refranes y locuciones en particular han llamado la atención de la comunidad científica también en la lengua española.

En sus albores, el estudio de las UF se basaba en una aproximación meramente abstracta, centrándose en su dimensión estructural y pragmática. Es a mediados de los años ochenta del siglo XX cuando, por primera vez, las investigaciones del ámbito hispánico se focalizan hacia el aspecto pragmático del universo fraseológico. Sin embargo, la complejidad de dichas unidades dificultó su estandarización y su encasillamiento en determinados parámetros y criterios, distinguiéndose y diferenciándose de otras UF para su posterior clasificación.

Como apunta Sevilla Muñoz (2012), a partir de la Edad Media se registran los primeros refraneros que, además de recopilar refranes o proverbios, recolectaron también las locuciones y las fórmulas rutinarias de la época. Entre los primeros ejemplos que se conocen, destacan los de Pedro Vallés (1549) y de Hernán Núñez (1555), antecedentes directos del interés por el estudio del habla en comunidades concretas que, posteriormente, evolucionó al área de conocimiento definido como fraseología.

La aproximación a la taxonomía de las UF es compleja, dada la variedad de clasificaciones existentes y los criterios que se han utilizado a lo largo

del tiempo. Desde un punto de vista diacrónico, la presentación de Corpas Pastor (1196:11) en su *Manual de fraseología* española no deja lugar a equivocaciones sobre los orígenes de dicha área de conocimiento: «la fraseología, como disciplina científica, se originó en la antigua URSS en la década de los cincuenta».

A continuación, se presentan las principales clasificaciones propuestas para la fraseología española, así como las aportaciones más relevantes que no se han concretizado en una clasificación holística, porque han tratado de manera sistemática solo algunas de las esferas del universo fraseológico.

2.3.1.1. Julio Casares

Es a mediados del siglo XX cuando el miembro de la Real Academia Española, Julio Casares, con su *Introducción a la lexicografía moderna* (1950) marca un hito en los estudios fraseológicos y paremiológicos. En esa época, el universo fraseológico era un objeto de estudio novedoso y, en gran medida, desconocido. Si bien la concepción del Casares sobre las unidades fraseológicas parte de una perspectiva meramente lexicográfica, algunas de las aportaciones que el autor granadino teoriza en su manual, y que se presentan a continuación, son de gran interés para esta tesis doctoral.

La primera hipótesis del lingüista es su original cuestionamiento de la libertad personal de los hablantes a la hora de producir actos de habla. Casares defiende por primera vez que los hablantes no son tan libres como creen, ya que no «construimos personalmente nuestras frases escogiendo aquí y allá materiales sueltos, como hace el albañil con los ladrillos», sino que «nos servimos en gran medida, como en las construcciones modernas, de armazones y de piezas 'prefabricadas'» (1950: 223), moldes que el hablante puede utilizar en función de su exigencia comunicativa. Sin embargo, como observa el autor (1950: 225),

no siempre es posible sustituir estas piezas prefabricadas ya que los elementos no siempre se prestan a nuevas combinaciones.

Desde la filosofía del lenguaje, dicha aproximación resulta particularmente interesante, ya que colisiona con la concepción del *libero arbitrio* a la hora de elegir palabras o términos que describen un concepto, demostrando en cambio que la morfo-sintaxis de un idioma, en buena medida, está constituida por esquemas, «armazones prefabricados» que se pueden montar y desmontar para cumplir con una necesidad comunicativa específica[21].

La clasificación de Casares promueve un acercamiento principalmente desde una perspectiva lexicográfica. De hecho, su intención es describir las dificultades que ciertas unidades presentan a la hora de insertarse en los diccionarios. Dicha orientación no le impide observar ciertas dificultades taxonómicas de las distintas unidades fraseológicas, y que reconoce que son caracterizadas por cierta inestabilidad.

Debido a su aproximación lexicográfica, Casares llega a separar dos áreas de la fraseología que hasta entonces se consideraban como un conjunto único[22] afirmando que los refranes no deben incluirse en un repertorio lexicográfico, salvo en casos muy excepcionales, por dos razones principales. En primer lugar, porque «el refrán no es un hecho expresivo reducible a una equivalencia conceptual unitaria; y, en segundo lugar, porque tampoco es una combinación verbal cuyos términos hayan adquirido significación especial al fundirse en un bloque» (1950: 197).

[21] La teoría de los esquemas fraseológico será reelaborada por Angela Mura (2012) en su tesis doctoral titulada «La fraseología del desacuerdo: los esquemas fraseológicos en español y en italiano», dirigida por Ruiz Gurillo.

[22] Dicha separación entre fraseología y paremiología será recuperada por autores como Ruiz Gurillo o Gerd Wotjak que se hicieron eco de esta concepción.

Con todo, Casares elabora un primer intento de clasificación definido «sintáctico-funcional». En este sentido, divide las unidades fraseológicas a partir de criterios categoriales, es decir, según la equivalencia que la UF tiene con las categorías gramaticales (sustantivos, verbo, adjetivo, etc.), y articula el universo fraseológico en cuatro pilares o tipos básicos: 1) locuciones; 2) frases proverbiales, 3) refranes y 4) modismos.

La clase de locuciones es la más esquematizada por el autor y la describe a través de una definición que sigue *grosso modo* repitiéndose, con modificaciones, en los más recientes manuales[23], es decir, una «combinación estable de dos o más términos, que funciona como elemento oracional y cuyo sentido unitario consabido no se justifica, sin más, como una suma del significado normal de los componentes» (1950: 170). La primera clasificación que articula para describir dicha clase de UF es distingue entre locuciones conexivas y significantes (Véase Figura 1). Como alude el término, las conexivas funcionan como nexo sintáctico en la oración, mientras que las significantes son las locuciones compuestas por lexemas llenos de significado. Por lo tanto, Casares divide las locuciones significantes en:

- Locuciones nominales
  a) Denominativas
  b) Singulares
  c) Infinitivas
- Locuciones adjetivales
- Locuciones verbales
- Locuciones participiales
- Locuciones adverbiales
- Locuciones pronominales

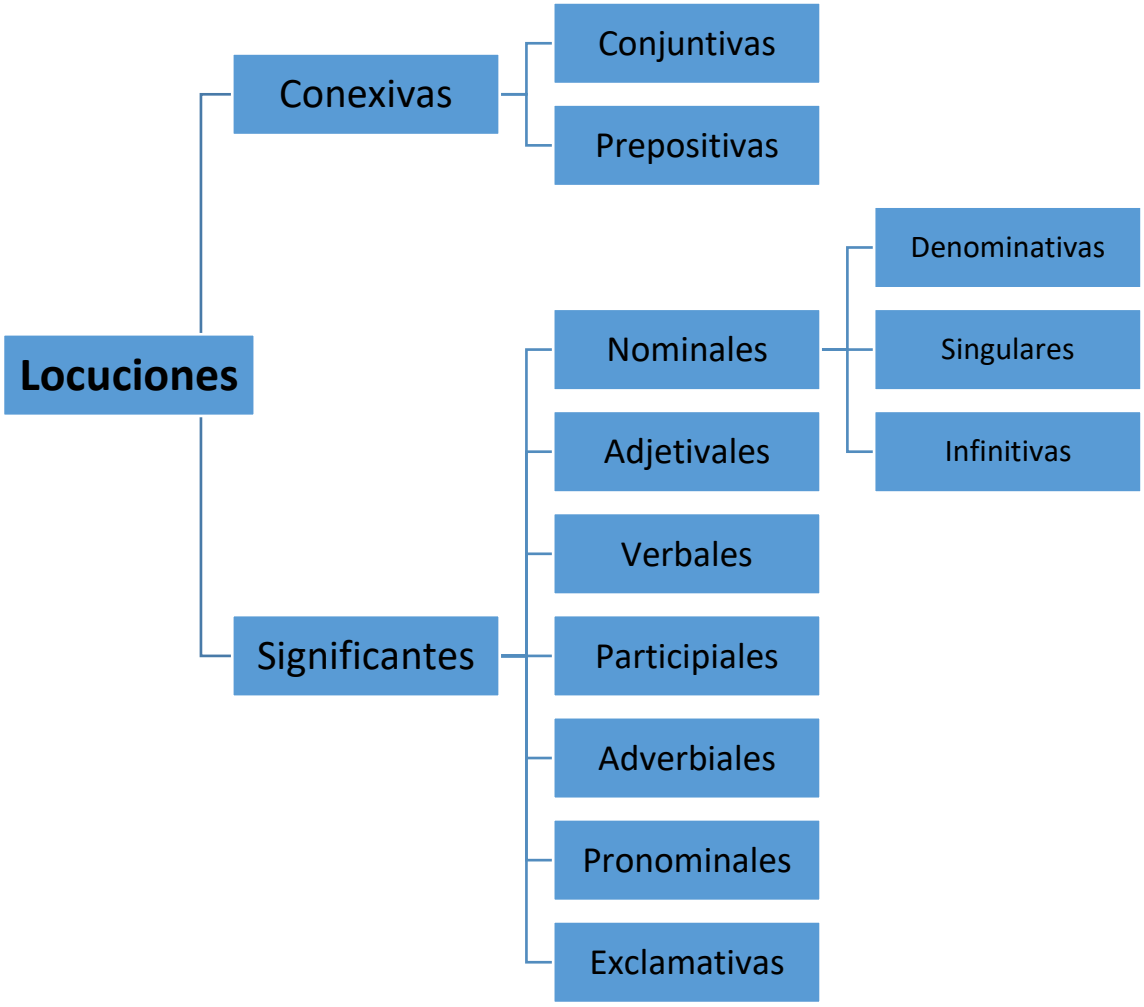
[23] Como se verá más adelante, tanto Zuluaga como Corpas Pastor se acercan en manera considerable a la definición elaborada por Casares a pesar de que entre una publicación y otra haya transcurrido casi medio siglo.

- Locuciones exclamativas

En lo que concierne a las locuciones adverbiales, Casares las distingue en función de la clase gramatical de los adverbios. Sin embargo, llama a la atención el hecho de que no considere los adverbios de modo. Corpas Pastor (1996), de hecho, considera que estas son las que más comúnmente definen las locuciones adverbiales, afirmando que «estas locuciones expresan distintos valores referenciales, especialmente modo [...]; cantidad [...]; localización en el tiempo [...]; y localización en el espacio» (1996: 100).

A diferencia de las demás, las clases de las locuciones participiales, es decir, cuyas locuciones contienen un participio, desaparecerán de la gran mayoría de las clasificaciones futuras. En este sentido, solo la cubana Carneado Moré (1985) mantendrá en su taxonomía las locuciones participiales como un grupo específico mientras que la tendencia más productiva que se ha registrado es agruparlas juntos con las locuciones adjetivales. Asimismo, Casares considera el grupo de las locuciones interrogativas como un tipo que ninguno de los estudiosos posteriores analizados en el presente trabajo contempla.

FIGURA 1. Clasificación de locuciones por Casares 1950.



Fuente: elaboración propia.

Dicha clasificación representa el punto de partida para los estudios fraseológicos españoles posteriores y, con sus debidas modificaciones y ampliaciones, será utilizada por autores como Zuluaga (1980) y Corpas Pastor (1996), hasta llegar a Mura (2015). A partir de ese momento, se despierta el interés de la comunidad de hispanistas en el mundo y, veinte años después, la Real Academia de la Lengua Española (RAE) ofrecerá una definición propia de locución en su Diccionario de uso de 1970.

Las frases proverbiales son la segunda tipología de UF que Casares identifica en su aproximación. El autor elige esta denominación frente a otras en boga en esa época porque «'expresión', 'giro' y 'frase' son vaguedades cautelosas que no nos sirven como términos técnicos» (1950:

185). Las frases proverbiales son, según Casares, parangonables a una unidad léxica autónoma[24] que contiene una tradición de ejemplaridad por consenso de una comunidad lingüística (1950: 188). Dichas frases en la lengua poseen la función de una cita cuyo valor expresivo, evocado a través de unidades léxicas, reside en el paralelismo que se establece entre el momento actual y el pasado. Su significado es, por lo tanto, idiomático y motivado por su origen, y a pesar de que su motivación pueda ser desconocida por su antigüedad, no supone limitación para que el hablante la utilice con frecuencia.

La tercera clase identificada por Casares son los refranes, definidos como:

> una frase completa e independiente, que en sentido directo o alegórico, y por lo general en forma sentenciosa y elíptica, expresa un pensamiento –hecho de experiencia, enseñanza, admonición, etc.– a manera de juicio, en el que se relacionan por lo menos dos ideas. En la mayoría de los casos las dos ideas están expresas. (Casares, 1950: 192)

Los refranes son descritos como frases completas e independientes en la oración que expresan un juicio personal relacionando dos o más ideas. Casares reconoce la dificultad de diferenciar netamente esta categoría de las frases proverbiales ya que, como se ha visto, estas también poseen dicha peculiaridad. Sin embargo, apunta algunas consideraciones que serán de gran ayuda para estudios futuros.

Los refranes describen un hecho, «una verdad valedera para toda la humanidad sin distinción de tiempos ni lugares» (Casares, 1950: 194), en cambio, las frases proverbiales son «privativas de los pueblos donde nacieron» (Casares, 1950: 196). Asimismo, los refranes apuntan al pasado, nacen en un tiempo alejado del presente y, por repetición, asumen

[24] Esta intuición resulta ser extremadamente novedosa en el discurso de Casares ya que, por primera vez, se llega a considerarlas como actos de habla autónomos que no necesitan otros elementos para cumplir su función de oración completa.

carácter proverbial. Sin embargo, también subraya la inutilidad e imposibilidad de reconocer los orígenes, ya que «muchos de los refranes han llegado hasta la actualidad vacíos de contenido, es decir, tuvieron su origen en un doble sentido, generalmente escatológico u obsceno, en una pulla encubierta o en una alusión a sucesos minúsculos justamente olvidados» (Casares, 1950: 202-203).

Finalmente, la última clase que identifica el autor son los modismos. Estos pueden ser frases proverbiales o locuciones pero no al contrario, es decir, «no todas las locuciones son modismos y que las denominativas no lo son, a nuestro entender en ningún caso» (Casares, 1950: 232). A pesar de que guarden una estrecha relación con dichas clases, lo modismos a su entender quedarían fuera de la lexicografía.

Casares enumera una serie de características que definen un modismo: la intraducibilidad[25], la fundamental anomalía gramatical, la inalterabilidad, y que la espontaneidad y los efectos cuantitativos que provocan que ciertos modismos sean impropios para la lengua escrita, mientras que otros sean fundamentales en la literatura, puesto que suponen «uno de los más eficaces recursos estilísticos» (Casares, 1950: 229). Finalmente, sobre la elección de la palabra modismo, el propio autor reconoce que el término no es exacto y que la lingüística no puede admitirlo, ya que contiene una vaguedad que no permite ceñir sus características.

Por todo ello se puede concluir que las aproximaciones de Casares a una clasificación del universo fraseológico del español representan un primer intento de esquematizar un área de la lingüística que paulatinamente será autónoma de la lexicografía. En este sentido, el autor se debe considerar como precursor para los fraseólogos de la lengua española y, en mayor o menor medida, de otros idiomas.

[25] Cabe aquí recordar que Casares conocía 18 idiomas y fue intérprete profesional de conferencias.

2.3.1.2. Alberto Zuluaga

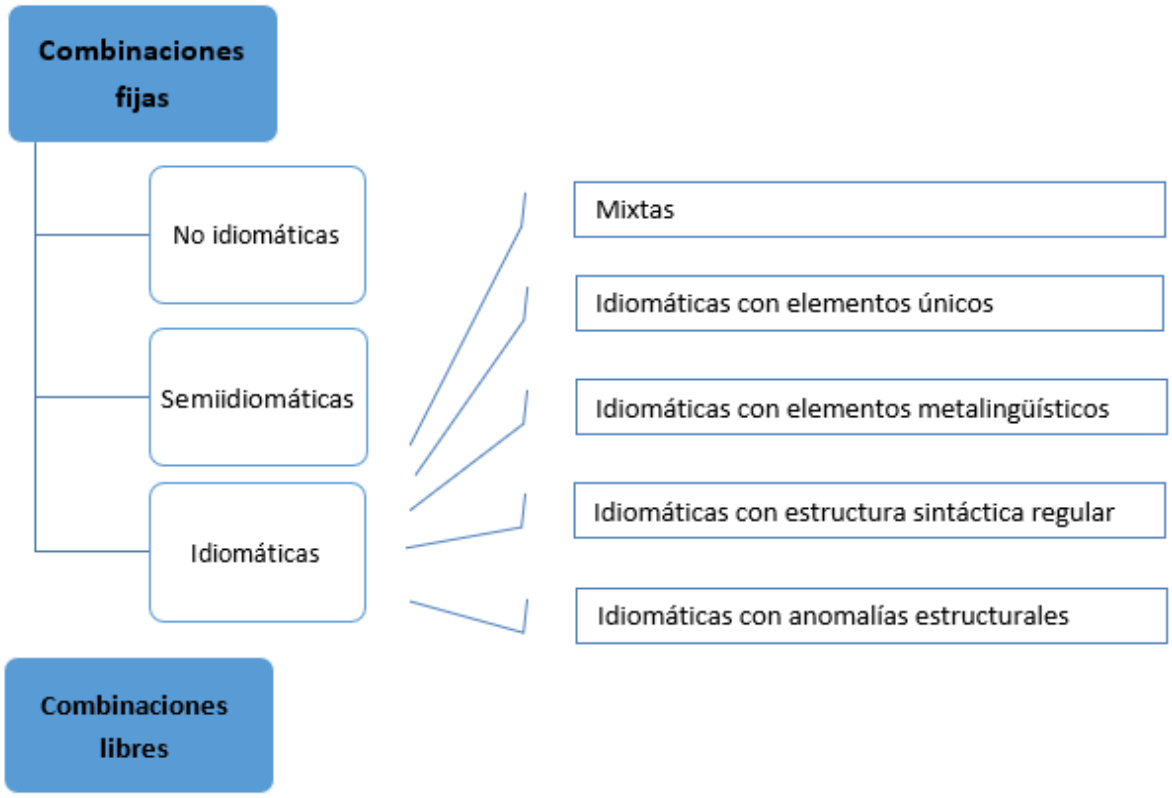
La siguiente clasificación relevante para el marco teórico del presente trabajo es la de Alberto Zuluaga. Su *Introducción al estudio de las unidades fijas* (1980) bebe de las intuiciones de Casares (1950), añadiendo algunas novedades relevantes que marcaron un hito para los estudios fraseológicos. En primer lugar, Zuluaga utiliza el término *expresiones fijas* o *unidades fraseológicas* para describir aquellas combinaciones compuestas, por lo menos, por dos palabras cuya relación intrínseca entre los componentes se caracteriza por la fijación y, en muchas de ellas, por la idiomaticidad. Por primera vez se emplea el término unidad fraseológica en la lingüística española.

Desde un punto de vista taxonómico, el autor realiza dos clasificaciones distintas de las unidades fraseológicas. La primera atiende a los rasgos de su estructura interna, es decir, la fijación y la idiomaticidad de las unidades; y, la segunda es elaborada en función de su valor semántico-funcional, puesto que las UF son empleadas como unidades en el discurso. De acuerdo con su estructura interna, Zuluaga argumenta que las unidades fraseológicas pueden ser 1) fijas y no idiomáticas como *dicho y hecho*, 2) semiidiomáticas como *lobo con piel de oveja* y 3) idiomáticas como *tirios y troyanos* (Zuluaga, 1980: 135-139).

Corpas Pastor critica parcialmente dicha clasificación y, en particular, destaca que las unidades que Zuluaga incluye entre las idiomáticas mezclan «criterios semánticos, sintácticos y metalingüísticos, sin homogeneidad alguna» (Corpas, 1996: 41). De hecho, las idiomáticas incluyen a su vez una serie de subtipos como las combinaciones y se corre el riesgo de sufrir cierta subjetividad, mezclando las unas con las otras. En este sentido, el lingüista distingue las combinaciones idiomáticas en cinco subtipos: 3.1) mixtas; 3.2) idiomáticas con elementos únicos; 3.3) idiomáticas con anomalías estructurales; 3.4) idiomáticas con elementos metalingüísticos y 3.5) idiomáticas con estructura sintáctica regular. En

contraposición con las unidades fraseológicas cuya estructura interna es fija, Zuluaga contrapone las combinaciones libres (Véase Figura 2).

Figura 2. Primera clasificación de Zuluaga según su estructura interna (1980).



Fuente: elaboración propia.

Con respecto a la segunda clasificación ofrecida por Zuluaga (1980: 141-165), es decir, en función del valor semántico-funcional de las unidades fraseológicas, el autor tiene en consideración las funciones semánticas que dichas unidades desempeñan en el discurso y elabora una taxonomía dividas en dos grandes grupos: 1) las locuciones y 2) los enunciados fraseológicos.

Las locuciones son definidas como expresiones fijas que necesitan combinarse con otros elementos de la frase para poder constituir un enunciado. En este sentido, no pueden existir de manera autónoma al ser

unidades inferiores al nivel de la oración. Las locuciones son a su vez divididas en dos subgrupos y cada uno de ellos a su vez en otros subgrupos. A saber:

1.1. locuciones como instrumentos gramaticales. Dicha categoría no presenta un significado léxico e incluye a su vez las:

1.1.1. Locuciones prepositivas.

1.1.2. Locuciones conjuntivas.

1.1.3. Locuciones elativas empleadas para intensificar los verbos, sustantivos y adjetivos.

1.2. locuciones como unidades léxicas. Estas incluyen a su vez las:

1.2.1. Locuciones nominales, funcionan como un sustantivo.

1.2.2. Locuciones adnominales, funcionan como un adjetivo.

1.2.3. Locuciones adverbiales, funcionan como componentes del predicado o determinando una oración.

1.2.4. Los esquemas fraseológicos.

1.2.5. Locuciones verbales, que equivalen a un verbo o a un predicado.

En cambio, los enunciados fraseológicos están formados por aquellas unidades independientes que pueden estar de manera autónoma en la frase en cuanto representan expresiones fijas con valor de enunciados completos. Zuluaga ofrece distintas clasificaciones para los enunciados fraseológicos. A continuación, se exponen con detalle los más relevante para el presente estudio (Zuluaga, 1980: 191-207). Los enunciados fraseológicos se subdividen a su vez en:

2.1. Fijos y de sentido literal.

2.2. Semiidiomáticos.

2.3. Idiomáticos. Esta categoría a su vez se subdivide en función de la estructura gramatical de sus componentes:

2.3.1. De una sola palabra;

2.3.2. Formados por un sintagma nominal;

2.3.3. Constituidos por una oración simple, nominal o verbal;

2.3.4. Formados por una oración compuesta;

2.3.5. Constituidos por una cita introducida con el verbo «decir» como *dijo la zorra a las uvas: no están maduras*;

2.3.6. Formados por una oración interrogativa unida a una respuesta como *¿Quién es tu amigo? Tu pariente en el mal*.

Zuluaga define los últimos cuatro subgrupos con el nombre de wellerismos o dialogismos, aunque se podrían también considerar refranes. Si bien no se observa una gran diferencia taxonómica entre la clasificación de Casares (1950) y la de Zuluaga, cabe destacar las aportaciones de gran relevancia y que superan la clasificación precedente.

En primer lugar, Zuluaga introduce dentro de las locuciones equivalentes a unidades gramaticales un nuevo tipo de locuciones que denomina *elativas*. Estas se caracterizan por actuar como elementos relacionales que intensifican los lexemas pertenecientes a las categorías del verbo, del nombre y del adjetivo (Zuluaga, 1980: 145-149). De este modo, la locución elativa *como un carretero* puede elativizar el verbo fumar. Corpas Pastor (1996), por su parte, critica dicha distinción realizada por Zuluaga al considerarla «además de innecesaria, poco acertada. […] Si la locución tiene estatus de palabra por definición (cf. Zuluaga, 1980: 23), aquella que no lo tenga no será, pues, una unidad fraseológica, es decir, no será una locución» (Corpas, 1996: 43), y añade que este carácter

enfático no es exclusivo de las locuciones por lo que considera más conveniente hablar del efecto intensificador de las UF en general.

Por otra parte, Zuluaga considera que Casares emplea de manera impropia la designación de locuciones nominales denominativas ya que el primero los considera simplemente como compuestos nominales. Con el objetivo de demostrar su afirmación, propone ejemplos de geminadas (papel moneda, falda pantalón, etc.) y de complejas (tocino de cielo, cabello de ángel, etc.) deduciendo que no son propiamente locuciones, sino compuestos nominales (Zuluaga, 1980: 57).

Como apunta Quiroga (2006), cabe destacar que la discusión entre compuestos y locuciones nominales ha sido siempre un tema de gran controversia. Zuluaga, apoyándose en el criterio de Benveniste (1966), considera que un compuesto es identificable por el hablante y posee una relación identificable y regular con el sentido total del compuesto. Por el contrario, si faltase dicha relación se trataría de una locución (Zuluaga, 1992: 10). Ante la falta de criterios adecuados para la distinción entre un compuesto y una locución, Corpas Pastor afirma que «hemos decidido considerar *compuestos* a todas aquellas unidades léxicas formadas por la unión gráfica (y acentual) de dos o más bases; y *locuciones*, a aquellas unidades que, presentando un grado semejante de cohesión interna, no muestran unión ortográfica» (Corpas, 1996: 93).

Por otra parte, las aportaciones de Zuluaga tienen el gran mérito de introducir en los estudios de la fraseología las fórmulas rutinarias y los refranes, aunar por primera vez, dos enfoques que existían en su época, y elaborar las clasificaciones previamente descritas en función de la fijación, la idiomaticidad y el valor semántico-funcional.

El trabajo de Zuluaga constituye el cimiento epistemológico del estudio de las unidades fraseológicas, siendo el primer autor en dedicar un monográfico a la fraseología española, puesto que Casares se ocupó de

ella solo en un capítulo de su *Introducción a la lexicografía moderna* (1950). La clasificación de Zuluaga, que toma como referencia la de Casares, ha alcanzado gran difusión en el ámbito hispánico y, en general, «ha sido la más aceptada hasta la publicación del estudio de Corpas Pastor» (Quiroga, 2006: 19), que aportará al estudio de la fraseología española importantes novedades y que se detallarán a continuación.

2.3.1.3. Gloria Corpas Pastor

En la actualidad, la clasificación del universo fraseológico Corpas Pastor (1996) «es la que mayor aceptación alcanza entre los estudiosos dedicados a la fraseología en España» (Quiroga, 2006: 45) y ha tenido gran repercusión también en Italia, como se explicará posteriormente. Para comprender de manera adecuada la taxonomía que la autora española realiza para ordenar el universo fraseológico, se deben referir las nociones coserianas, a su vez inspiradas en Sausurre, de norma, sistema y habla. Dicha tricotomía inspiró a Thun (1978) para ofrecer una clasificación de las UF que a su vez promovieron la elaboración de la taxonomía de Corpas Pastor. Coseriu (1962) considera que la dicotomía lengua – habla sausurreana es insuficiente:

> La primera insuficiencia está en la identificación inicial entre individual y concreto social y formal (funcional. […] La segunda insuficiencia de la dicotomía sausurreana, en su foma última, es la de ser demasiado rígida, es decir, de ignorar el punto en que «lengua» y «habla» se encuentran y se combinan, o sea, el «acto verbal». […] En tercer lugar, la dicotomía sausurreana aparece demasiado rígida también por la concepción del individuo que tiene de Saussure, un individuo completamente separado de la sociedad y que no sería él mismo «colectividad». (Coseriu, 1962: 30-32)

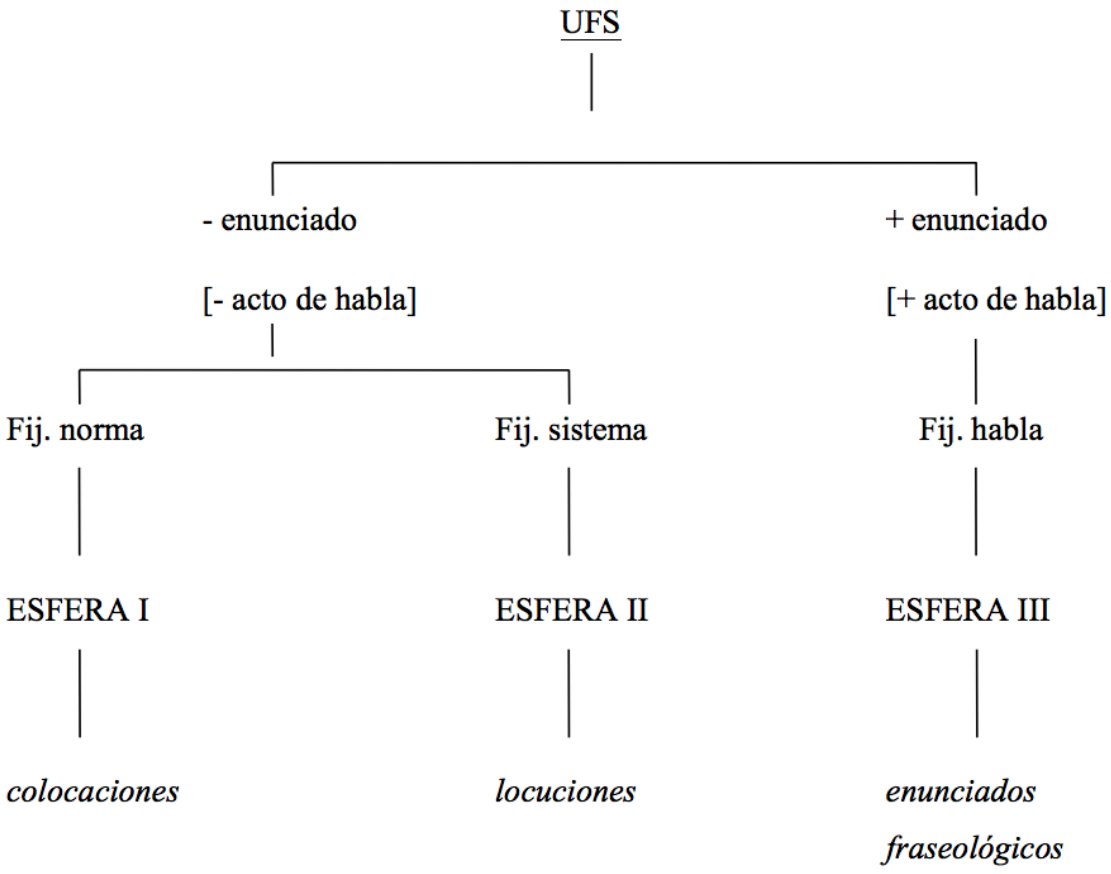
Dicha dicotomía es demasiado rígida para Coseriu porque no permite comprender cómo se pasa del sistema abstracto a las realizaciones concretas del habla. De este modo, el autor introduce un concepto intermedio entre el *sistema* y el *habla*: el concepto de *norma*, refiriéndose a las reglas que se fijan social o tradicionalmente en el habla o, según sus propias palabras: al «conjunto formalizado de las realizaciones tradicionales del sistema» (Coseriu, 1981: 316-317). En este sentido, la norma reduce las posibilidades del sistema al ser realizaciones lingüísticas tradicionales. Se trata de realizaciones obligatorias, dependientes del entorno geográfico-lingüístico y sociocultural, que cambian en función de la comunidad de hablantes. Es, por ende, el uso más correcto de una lengua y aceptado por una determinada comunidad de hablantes.

Corpas Pastor (1996) combina el criterio de fijación (en la norma, en el sistema o en el habla) con el de enunciado. De acuerdo con Zuluaga (1980), la autora se refiere con enunciado a «una unidad de comunicación mínima producto de un acto de habla, que corresponde generalmente a una oración simple o compuesta pero que también puede constar de un sintagma o una palabra» (1996: 51).

En un primer nivel, Corpas establece una primera distinción entre las colocaciones y las locuciones que no constituyen un acto de habla, es decir, un enunciado completo *per se*, y las que sí lo son, los enunciados fraseológicos (véase Figura 3). El tipo de fijación que presentan estas unidades fraseológicas permite diferenciar tres grandes subgrupos en un segundo nivel, que denomina Esferas (Corpas, 1996: 50). La primera esfera está compuesta por las colocaciones, UF fijadas en la norma y que no son actos de hablas completos; la segunda esfera incluye las *locuciones*, fijadas en el sistema y que tampoco son actos de habla; finalmente, la tercera esfera está caracterizada por los *enunciados fraseológicos*, que, al pertenecer «exclusivamente al acervo socio-cultural de la comunidad hablante» (Corpas, 1996: 51), están fijadas y constituyen actos de habla

realizados por enunciados completos. Cada esfera es susceptible de ser dividida en distintos subtipos, de acuerdo con una serie de criterios complementarios. A continuación, se propone la clasificación gráfica presentada por la autora en su manual.

Figura 3. Primer nivel de clasificación propuesto por Corpas Pastor (1996)



Fuente: Corpas Pastor (1996: 52).

Según Corpas Pastor, las colocaciones que componen la primera esfera son unidades fraseológicas a todo derecho. Estas se generan a partir de reglas (la norma) pero presentan cierto grado de restricción determinada por el uso. Son unidades prefabricadas que no pueden alterarse como *Estallar una guerra; Momento crucial; Manada de lobos;* etc., y que la autora define, basándose en la definición de Haensch *et al.* (1982: 251), como «aquella propiedad de las lenguas por la que los hablantes tienden

a producir ciertas combinaciones de palabras entre una gran cantidad de combinaciones teóricamente posibles» (1996: 82).

En función de la categoría gramatical y de la relación sintáctica entre los colocados, la autora elabora una taxonomía de dicha primera esfera de las colocaciones. Corpas Pastor incluye en su clasificación seis son los tipos de colocaciones: 1) sustantivo (sujeto) + verbo; 2) verbo + sustantivo; 3) adjetivo + sustantivo; 4) sustantivo + preposición + sustantivo; 5) verbo + adverbio; 6) adjetivo + adverbio. De ellas, las más frecuentes son la segunda y la tercera, según la autora.

La segunda esfera, las locuciones, está compuesta por unidades que están fijadas en el sistema y que no componen un enunciado completo. Dichas unidades fraseológicas están caracterizadas por su «institucionalidad, su estabilidad sintáctico-semántica y su función denominativa» (Corpas, 1996: 89), características que las distinguen de las combinaciones libres de palabras. Como toda unidad fraseológica, las locuciones están compuestas por al menos por dos términos que no muestran unión ortográfica, de otra manera se trataría de palabras compuestas, muy propias de las lenguas aglutinantes.

Para reconocer que nos encontramos frente a una locución, la autora explica que es posible utilizar tres pruebas para comprobar la estabilidad de dichas unidades. Si una de las tres pruebas concluye con una nueva secuencia gramaticalmente correcta en el plano del sistema pero sin cohesión semántica, nos encontraríamos, sin duda, frente a una locución. Los tres métodos propuestos son los siguientes: 1) sustitución de uno de los elementos constituyente por otro término (sinónimo, hiperónimo, hipónimo); 2) eliminación de uno de los constituyentes, o adiciones; 3) deficiencias transformativas: como la posibilidad de invertir el orden de sus componentes o del artículo. Si el orden de las palabras o los cambios se producen sin alterar la cohesión semántica de la unidad, se trataría pues de una combinación libre de palabras.

La taxonomía que elabora Corpas Pastor para las locuciones toma como referencia las elaboradas por Casares, es decir, ambos se basan en la función oracional que estas desempeñan. Por otra parte, a pesar de que la autora critica parte de esta taxonomía, recupera parcialmente la clasificación de locuciones de Zuluaga, a la que añade una nueva función oracional, las locuciones clausales. Así distingue entre siete grupos de locuciones:

1) locuciones nominales, cuyos patrones más productivos son formados por sustantivo + adjetivo y sustantivo + preposición + sustantivo; otro patrón frecuente es, según la autora, el formado por sustantivo + sustantivo; finalmente, entre las locuciones nominales tienen cabida las que Casares ha definido como locuciones infinitivas, las cláusulas sustantivadas y, por último, las expresiones deícticas carentes de otro significado léxico[26].

2) Locuciones adjetivas que, según insiste la autora, «desempeñan las funciones oracionales básicas de atribución y de predicación» (Corpas, 1996: 97), están principalmente compuestas por adjetivo/participio + preposición + sustantivo y también por adjetivo + conjunción + adjetivo; también se incluyen dentro de las locuciones adjetivas las comparaciones estereotipadas y las locuciones encabezadas por un relativo. Finalmente, la autora distingue otro grupo de locuciones adjetivas formadas por un sintagma prepositivo.

3) Locuciones adverbiales, que en su gran mayoría son sintagmas prepositivos que poseen cierta relación con la categoría gramatical del adverbio al desempeñar sus mismas funciones semánticas. Estas unidades expresan distintos valores referenciales como modo, cantidad, tiempo y espacio y se trataría de una categoría bastante abundante en la que tienen cabida distintos tipos de locuciones adverbiales como los

[26] Los elementos de estas unidades fraseológicas no tienen significados léxicos.

modificadores oracionales[27], los sintagmas cuyo núcleo es un adverbio e, incluso, los sintagmas sustantivos.

4) Locuciones verbales, que «expresan procesos, formando los predicados, con o sin complementos» (Corpas, 1996: 102). Corpas Pastor identifica distintos patrones que caracterizarían dicha categoría: verbo + conjunción + verbo (+ complemento); verbo + pronombre (+ partícula); verbo + partícula (+ complementación); verbo copulativo + atributo; verbo + complemento circunstancial; y verbo + suplemento; verbo + objeto directo con complementación opcional. Muchas locuciones verbales pueden formar por sí mismas los predicados de oraciones. Otras, en cambio, contienen casillas vacías a las cuales se debe insertar un elemento. Finalmente, Corpas Pastor recuerda que existe un gran número de locuciones verbales que están fijadas en su forma negativa[28].

5) Locuciones prepositivas son aquellas unidades susceptibles de formar una unidad de tipo prepositivo. Están formadas por un adverbio o un sustantivo adverbializado + una preposición. Además, la profesora afirma que existen dos tipos de unidades: las que están constituidas por adverbios que funcionan solos y las que siempre necesitan un adyacente especificado.

6) Las locuciones conjuntivas se diferencian de las demás porque no forman sintagmas por sí mismas. La autora las clasifica según su función, es decir, las coordinantes y las subordinantes. Las primeras engloban las locuciones conjuntivas distributivas o disyuntivas y las adversativas, mientras que las subordinantes suelen, en cambio, introducir cláusulas finitas. Los valores que pueden presentar estas locuciones son los

[27] Con el fin de recordar que los límites de una categoría a otra son más bien líquidos (para utilizar una terminología *baumanhiana*) que fij, entre ellos la autora pone el ejemplo de la UF *según mi* y que, a nuestro entender, podría ser clasificado de igual manera como una fórmula rutinaria psico-social.

[28] García-Page (2008) reafirma dicha característica de las locuciones.

condicionales; concesivos; causales; consecutivos; finales; modales; comparativos y temporales.

7) Las locuciones clausales concluyen la taxonomía de Corpas Pastor de las locuciones. La autora reconoce que este grupo engloba a un gran número de locuciones «de diversa índole formados por varios sintagmas de los cuales al menos uno de ellos es verbal» (Corpas, 1996: 109) y que funcionan como elementos oracionales. Corpas Pastor distingue dos tipos de locuciones clausales: Al primer tipo pertenecen aquellas que poseen una sola casilla vacía que corresponde al objeto o al complemento de la expresión; mientras que en el segundo grupo se enmarcan aquellas locuciones formadas por cláusulas enteras, pero que no poseen independencia textual ni fuerza ilocucionaria[29]. Dicho tipo de locuciones, como se verá más adelante, representa una importante novedad en comparación con las clasificaciones presentadas hasta el momento.

La tercera y última esfera a las que hace referencia Corpas Pastor en su manual está formada por los enunciados fraseológicos[30]. Como ya se ha explicado, estas unidades fraseológicas, a diferencia de las colocaciones y las locuciones, sí representan actos de habla por sí mismas y presentan fijación interna y externa. En los enunciados fraseológicos la autora incluye dos subgrupos: las paremias y las fórmulas rutinarias, insistiendo en que la separación entre una y otra no es neta, y que incluso a veces resulta difícil distinguirlas de las locuciones.

Las paremias[31] representan el primer tipo de enunciados fraseológicos y funcionan como secuencias autónomas que disponen de una entonación propia, lo que las convierte en unidades de comunicación

[29] Para la definición de fuerza ilocucionaria entendida como el conjunto de fuerzas que permiten la realización de un acto o que llevan a cabo el acto de decir algo, véase Austin (1996).

[30] Dicha nomenclatura no es nueva a los estudios fraseológicos. Como se ha explicado, ya Zuluaga (1980: 192) las definía como unidades con «secuencias autónomas de habla […] unidades de comunicación mínima».

[31] Hasta la clasificación de Corpas Pastor, las paremias, igual que las colocaciones, no se analizaron como parte integrante del universo fraseológico.

mínimas. Por lo tanto, las paremias se distancian de las fórmulas rutinarias (segundo tipo de la tercera esfera) en cuanto poseen un significado referencial, mientras que el de las fórmulas rutinarias es de tipo social, expresivo y discursivo. Además, si las paremias presentan autonomía textual, las fórmulas rutinarias se supeditan a situaciones o circunstancias extralingüísticas concretas.

Inspirándose en la tesis de Arnaud (1991), la autora distingue dentro de las paremias entre enunciados de valor específicos, es decir, UF que no tienen valor de verdad general pero que constituyen enunciados fraseológicos textuales; y citas, verdades generales que, a diferencia de los refranes, se conoce su origen que normalmente provienen de la literatura, frases de personajes famosos de alguna época o, a menudo, de la biblia. Según la autora, y de acuerdo con Casares (1950), el refrán no siempre posee orígenes conocidos. Sin embargo, Corpas Pastor afirma que «el refrán es la paremia por excelencia, pues en él se dan las características definitorias y definidas por Arnaud (1991): lexicalización, autonomía sintáctica y textual, valor de verdad general y carácter anónimo».

Las fórmulas rutinarias representan el segundo tipo de los enunciados fraseológicos pertenecientes a la tercera esfera. También llamadas por otros autores frases idiomáticas pragmáticas o fórmulas comunicativas. Estas se dividen a su vez en fórmulas discursivas y fórmulas psicosociales. Las fórmulas discursivas, introducidas en la fraseología española por Zuluaga, son secuencias memorizadas y lexicalizadas que dan dirección al discurso, varían en función de la situación comunicativa y muestran la actitud del emisor hacia lo que dice. Estas comprenden las fórmulas de apertura, cierre y de transición.

Las fórmulas psicosociales, en cambio, «desempeñan funciones facilitadoras del desarrollo normal de la interacción social, o bien funciones de expresión del estado mental y los sentimientos del emisor»

(Corpas, 1996:192). Corpas Pastor observa que los criterios de clasificación no son semánticos sino pragmáticos. Por lo tanto, los subgrupos que conforman las fórmulas psicosociales se rigen en función de la carga ilocucionaria y los tipos de actos de habla que realizan las unidades fraseológicas y son los siguientes: fórmulas expresivas, comisivas, directivas, asertivas, rituales y misceláneas.

La clasificación elaborada por Corpas Pastor en su *Manual de fraseología española* sigue siendo el modelo que mayor aceptación ha recibido entre los estudios fraseológicos y supone un gran avance para el análisis de las UF por varias razones. En primer lugar, las colocaciones asumen un papel primario. Si bien con Casares y Zuluaga las colocaciones no tenían cabida en el universo fraseológico y, en general, dichas unidades no habían sido tratadas con la atención que merecían, en la actualidad, las colocaciones reciben gran atención constituyendo objeto de investigación por parte de los especialistas. En este sentido, se puede afirmar que se cubre un vacío existente en fraseología hasta dicha clasificación.

Cabe destacar que autores como Penadés (1997) afirman que las colocaciones no deberían de considerarse parte del universo fraseológico dado que las solidaridades léxicas[32] se diferencian de las unidades fraseológicas en cuanto las primeras no forman en sí mismas ni por sí mismas una estructura lexemática. Zuluaga comparte la idea de que las colocaciones no deberían formar parte de un sistema fraseológico, puesto que no poseen alguna característica especial ni con respecto al material léxico, ni en cuanto al contenido semántico, sino que

[32] Concepción elaborada por Eugenio Coseriu (1952), y posteriormente reelaborada por Zuluaga (1980), similar al fenómeno que define las colocaciones. Se trata de una relación sintagmática en la que existe una clase determinante y una clase determinada siendo esta última la que contiene un rasgo distintivo de la clase: el archilexema o el lexema de la clase determinante. Las solidaridades léxicas pueden ser de tres tipos y la que más afinidades tiene con la noción de colocación sería la solidaridad léxica multilateral. Para más información sobre este concepto véase Muñoz Núñez, Mª Dolores (2011): «Reflexiones sobre el concepto de solidaridades léxicas de E. Coseriu en el marco de los recientes estudios sobre restricciones léxicas y colocaciones», en Lorenzo Hervás, 20, 109-121.

su limitación es debida exclusivamente a su uso frecuente (Zuluaga, 1997: 22).

Otro aspecto novedoso de la clasificación de Corpas Pastor tiene que ver con las locuciones que son estandarizadas según su función oracional. En este sentido, la autora añade una nueva clase, las locuciones clausales, término probablemente influenciado por las locuciones adverbiales cláusula de Zuluaga (1980: 139), definidas por el autor como aquellas que modifican o comentan el contenido total de una oración. Asimismo, cabe destacar que hay autores como Castillo (2002), que rechazan la inclusión de estas estructuras en un subgrupo especial y las incluyen como parte de las locuciones verbales.

A pesar de que las similitudes con Zuluaga guarden una relación muy estrecha con la tercera esfera, «Corpas reordena y reagrupa las ideas» (Quiroga, 2006: 25). De hecho, la autora reúne en un solo grupo las fórmulas de fijación pragmática de Zuluaga y las esquematiza en un modelo que, hasta ahora, no ha sido refutado.

La clasificación que Corpas Pastor (1996) propone reúne todas las unidades que se considera que deberían componer el universo fraseológico entendido como un ámbito complejo en el que, a pesar de su taxonomía estable no posee límites rígidos sino zonas en las que las unidades fraseológicas, se entremezclan debido a características que complican su encasillamiento en una determinada esfera. Además, cabe destacar que dicho modelo ha sido aplicado con éxito en otros idiomas como el catalán y el valenciano, por Josep Guia (1999: *Fraseologia i stil. Enigmes literaris a la Vàlencia del segle XV*. Barcelona: Tres I Cuatre) y por P. Sancho Cremades (1999: *Introducció a la fraseología. Aplicació al valencià colloquial*, Valencia: Denes); al gallego, en el centro Ramón Piñeiro; al euskera[33] y, se realizó una aproximación a

[33] véanse los artículos de Pello Esnal y Xabier Kalzakorta, publicados en Euskera, 2001, nº 1, la revista de la Academia de la Lengua Vasca.

la lengua italiana llevada a cabo por Jorge Jesús Leiva Rojo (2005: *La recepción de la obra de Rossana Campo en España y Alemania: la traducción de los componentes oral y fraseológico de Mai sentita cosi bene*, Tesis doctoral); al polaco (2006: *Fraseología comparada del polaco y del español: su tratamiento en los diccionarios bilingües*, Tesis doctoral), etc.

2.3.1.4. Leonor Ruiz Gurillo

La siguiente clasificación del universo fraseológico que se presenta por sus interesantes aportaciones es la de Ruiz Gurillo. En su libro *Aspectos de fraseología teórica española* (Gurillo, 1997), la autora articula su clasificación a través de los conceptos de centro y periferia dentro del sistema de la lengua propuesto por la Nueva Escuela de Praga[34]. Uno de los lingüistas que más ha dedicado sus estudios a la dicotomía entre centro y periferia es Vachek, quien explica que:

> It can be demonstrated that also within the system of language there can be, more or less clearly, distinguished two areas: the centre of the language system in which the mutual relations linking up the elements of the system set out very clearly, and, on the contrary, the periphery of the system, in which the ties linking its elements to those of the centre appears rather loose and sometimes may even be on the point of vanishing altogether. Further, in the periphery of the language system may appear as rather vague those relations and those distinctions which in the centre of the system appear to be clearcut and separated from one another by distinctly drawn lines[35]. (Vachek, 1976: 25)

[34] En 1964 aparece la revista *Travaux Linguistiques de Prague* y, en su segundo número, se abordará dicho tema, fundamental en los estudios lingüísticos checos a partir de los años '60 del siglo XX.

[35] «Se puede demostrar que también dentro del sistema lingüístico se pueden distinguir, más o menos claramente, dos áreas: el centro del sistema lingüístico en el que las relaciones mutuas que unen los elementos del sistema se establecen muy claramente y, por otra parte, la periferia del sistema, en la que los lazos que unen sus elementos a los del centro aparecen más relajados y a veces incluso pueden estar a punto de desaparecer por completo. Asimismo, en la periferia del sistema lingüístico las relaciones

Las zonas periféricas representan un área dinámica más propensa al cambio, y esta transformación puede realizarse en direcciones divergentes ya que puede producirse una atracción de los elementos periféricos hacia el centro del sistema o, por el contrario, pueden alejarse hasta desaparecer. Ruiz Gurillo, reconociendo la dificultad de enmarcar las UF en casillas específicas y establecer límites entre dichas unidades, ya que no todas ellas responden a todas las características que las definen, traslada este modelo a la fraseología reconociendo que «la visión en centro y periferia permite dicha estructuración gradual» (Ruiz Gurillo,1997: 73).

La autora considera que las clases nucleares, es decir, las que presentan mayor fijación e idiomaticidad, no poseen una variación significativa como les sucede a las clases periféricas, que reflejan una gran riqueza. Establece, por lo tanto, una escala gradual que va de las locuciones totalmente fijas e idiomáticas hacia aquellas con un alto grado de idiomaticidad. Entre ellas, las combinaciones semifiijas. La progresión de continuidad en las concepciones estrechas viene establecida fundamentalmente por la fijación y la idiomaticidad: cuanto mayor sea la idiomaticidad de la UF, mayor será su fijación y, por lo tanto, mayor el acercamiento al núcleo.

En cambio, en las zonas periféricas que confinan con las unidades libres se encuentran tanto las unidades sintagmáticas, como las combinaciones de palabras frecuentes (las colocaciones), que la autora excluye de su clasificación ya que, en su opinión, no presentan fijación sino una simple «afinidad» entre sus componentes (Ruiz Gurillo, 1997: 7). Según esta aproximación teórica, en el centro del sistema los elementos están fuertemente interrelacionados, mientras que en su periferia se

pueden ser vagas y las distinciones que en el centro del sistema parecen estar claras y separadas las unas de las otras por líneas claramente trazadas» Traducción propia.

encuentran los elementos que no están tan estrechamente vinculados en el sistema.

La autora se adscribe a la visión estrecha de la fraseología, como hará García-Page (2008), centrando su taxonomía en aquellas unidades equivalentes a una palabra o un sintagma. Ruiz Gurillo afirma que se trata de una «progresión que va desde la irregularidad a la regularidad de la improductividad a la productividad, de la fraseología a la sintaxis, pero ante todo constituye una diferenciación artificial que responde a un *continuum* de difícil segmentación» (Ruiz Gurillo, 1997: 122).

Con el objetivo de llegar a una clasificación del universo fraseológico, la autora procede a una diferenciación en función de cinco niveles: fonético-fonológico, morfológico, sintáctico, léxico-semántico y pragmático. Tras describir los criterios que definen una UF, Ruiz Gurillo propone su taxonomía y parte de un primer criterio formal de tipo distribucional identificando tres clases de sintagmas fraseológicos: los nominales, los verbales y los prepositivos. Finalmente, por cada uno de ellos, establece su centro y su periferia. A continuación, se resume la taxonomía propuesta:

Sintagmas nominales fraseológicos:

- Locuciones totalmente fijas e idiomáticas con palabras diacríticas y/o anomalías estructurales: *agua de borrajas*.
- Locuciones idiomáticas en diversos grados: *caballo de batalla*.
- Locuciones mixtas: *dinero negro*.
- Colocaciones: *agua de colonia*.

Sintagmas verbales fraseológicos:

- Locuciones con palabras diacríticas: *tomar las de Villadiego*.
- Locuciones totalmente fijas e idiomáticas: *dorar la píldora*.
- Locuciones semiidiomáticas: *echar raíces*.

- Locuciones mixtas: *vivir del cuento*.
- Locuciones meramente fijas: *correr mundo*.
- Locuciones con variantes: *no importar un pimiento! un bledo*.
- Unidades sintagmáticas verbales: *hacer uso, tomar un baño*.
- Otras colocaciones: *guiñar un ojo*.

Sintagmas prepositivos fraseológicos:

- Locuciones totalmente fijas e idiomáticas con palabras diacríticas y/o anomalías estructurales: *a la virulé*.
- Locuciones totalmente fijas e idiomáticas: *a menudo*.
- Locuciones parcialmente fijas e idiomáticas: *a mano*.
- Locuciones meramente fijas: *en público*.
- Locuciones con variantes: *de (muy) buen grado*.
- Locuciones con casillas vacías: *a mí (tu, su...) juicio*.
- Creaciones locucionales analógicas: *a gritos, a golpes*.
- Esquemas fraseológicos: *cara a cara*.

Cabe destacar que la clasificación de Leonor Ruiz Gurillo se distancia solo parcialmente de la que elabora Corpas Pastor (1996), por dos razones principales: en primer lugar, al basarse, como se ha explicado con anterioridad, en la concepción estrecha de la fraseología, excluye la esfera I de las colocaciones y la esfera III compuesta por los enunciados fraseológicos[36], y, en segundo lugar, por añadir las que denomina como locuciones marcadoras, introducidas como rasgo fundamental en la cohesión del discurso y que no funcionan de forma independiente sino son parte de la sintaxis oracional[37].

Otro trabajo interesante de la autora para el marco teórico de la presente investigación es el monográfico *Las locuciones en español actual* (2001) en el que, a través de un enfoque didáctico, propone una

[36] Temas que afrontará solo posteriormente en Ruiz Gurillo, (2000).

[37] Para ampliar el tipo de locuciones propuesto por Ruiz Gurillo (2010).

clasificación de las locuciones inspirada en Casares (1950), Zuluaga (1980) y Corpas Pastor (1996). A pesar de que en su introducción afirma que «el enfoque y el planteamiento inicial lo convierte en un material útil para estudiantes, profesores y público en general, interesado por estas unidades complejas de la lengua» (Ruiz Gurillo, 2001: 9), la estructura del manual y el planteamiento de los ejercicios propuestos en la parte final del trabajo parecen estar más bien dirigidos a un público especializado[38].

La autora se inspira en las tres clasificaciones que la preceden, descritas en los apartados precedentes, afirmando que «hasta el momento [...] la clasificación de locuciones más adecuada reflejaría la propuesta de Corpas Pastor» (Ruiz Gurillo, 2001: 45). Ruiz Gurillo identifica siete clases de locuciones sustituyendo las locuciones conjuntivas identificadas por Corpas Pastor (1996) con una nueva clase que denomina *Locuciones marcadoras.*

Las funciones de las locuciones marcadoras son, según la autora, de gran utilidad, puesto que «contribuyen a la cohesión de un texto, facilitan la progresión argumentativa o matizan enunciados o guían las inferencias que se han de extraer del discurso» (Ruiz Gurillo, 2001: 46). En este sentido, este tipo de locuciones, que definiría más bien una clase pragmática de locuciones, incluye locuciones como *sin embargo, no obstante, de todas maneras, en el fondo, por supuesto,* etc., siendo estas locuciones empleadas para enlazar, ordenar, rectificar, aclarar, puntualizar o ejemplificar enunciados y pueden tener valores opositivos, consecutivos, temporales, etc. (Ruiz Gurillo, 2001: 56-58).

[38] Los primeros cuatro capítulos están dedicados a ofrecer un marco teórico de la fraseología. En este sentido, la autora elabora una introducción en la que desarrolla su objeto de estudio, la fraseología, y, a continuación, define los rasgos específicos de esta ciencia, que identifica en la fijación, la idiomaticidad y las relaciones entre estas dos propiedades de las UF. En el tercer capítulo se aborda definición de las locuciones y la relación que estas tienen con otras esferas de la fraseología como las colocaciones, las metáforas y las unidades sintagmáticas. Finalmente, propone una nueva clasificación específica para las locuciones españolas.

Ruiz Gurillo aporta a la clasificación del universo fraseológico una propuesta en la que la gran novedad está representada por una nueva clase de locución que denomina como marcadoras y en la que inserta numerosas locuciones de distinto tipo (adverbiales, adjetivales, preposicionales, etc.). Algunas sirven para la intensificación semántica, como *hasta el moño* o *que no veas*; algunas son elementos sintácticamente suprimibles pero necesarios en un nivel pragmático, como *de alguna manera, en cierto modo* o *en principio*; otras no son suprimibles. Sin embargo, el hecho de que sea una miscelánea hace complicada la labor de catalogación de esta clase de locuciones marcadoras.

A pesar de la intención didáctica manifestada en la introducción de su manual, el espacio dedicado a la didáctica de la fraseología es más bien reducido. En particular, Ruiz Gurillo dedica un breve apartado sobre las relaciones entre enseñanza y aprendizaje y sobre el dominio lingüístico en términos de competencia y conocimiento de las UF. La autora defiende que ciertas unidades, como las fórmulas, se suelen y se debería enseñar/aprender en los niveles básicos, mientras que las locuciones y las paremias solo a partir de los niveles intermedios y avanzados, puesto que es cuando el alumnado ya ha adquirido el núcleo estructural de la LE en cuestión (Ruiz Gurillo, 2001: 89).

Finalmente, la autora propone una serie de ejercicios (Ruiz Gurillo, 2001: 97-100) para poner en práctica las bases teóricas ofrecidas en los capítulos precedentes. No se analizan en este trabajo doctoral porque se considera que ejercitaciones del tipo: «en los siguientes pares, diferencie la locución más motivada de la menos motivada» o «en los siguientes pares, separe las metáforas, semilexicadas o libres, de las locuciones» están dirigidas sobre todo a personal especializado en fraseología más que al alumnado y, por ello, excede el objeto de estudio de la presente tesis que se centra en el aprendiz como fin último de la fraseología. De

hecho, se considera que «es primordial acercar los resultados de estos trabajos a la realidad del aula y a los profesores de lenguas, con el fin de que se incorporen a su actividad docente» (Olimpo de Oliveira Silva, 2004: 133).

2.3.1.5. Mario García-Page Sánchez

A continuación se presenta la clasificación propuesta por Mario García-Page Sánchez en su *Introducción a la fraseología española* (2008). Como observa Antonio Pamies Bertrán, en su reseña para la revista *Language Design*:

> No había salido un manual de fraseología desde el célebre de Gloria Corpas (1996), por lo que, tras doce años de gran producción científica española en este campo, era necesario hacer balance crítico de los progresos e identificar de nuevo las carencias e incógnitas de una especialidad que se ha enriquecido mucho -pero también se ha complicado- durante dicho período. (Pamies Bertrán, 2008: 150-151).

A diferencia de la clasificación de Corpas Pastor, el autor reduce el universo fraseológico en su clasificación. Como Ruiz Gurillo, García-Page se adscribe a la visión estrecha de la fraseología, es decir, considera que las únicas UF que entrarían a todo derecho en el universo fraseológico son las locuciones. Con respecto a las colocaciones, el autor las plantea como un fenómeno propio de la sintaxis, excluyéndolas de su clasificación. En lo que concierne a las paremias y los refranes, García-Page en línea con otros autores (Casares 1950; Zuluaga, 1980), opina que estas son objeto de estudio de la paremiología, por lo que tampoco tendrían un lugar en los estudios de fraseología.

Para elaborar su taxonomía, García-Page parte de las aportaciones de Casares, que compara con las de otros autores (García-Page, 2008: 91-92). Su clasificación coincide en gran medida con la esfera II de Corpas

Pastor (1996), aunque con algunas diferencias que se explican a continuación. El autor clasifica las locuciones en ocho categorías según su función oracional:

1. Locuciones nominales, en las que incluye las que Casares define como infinitivas y, posteriormente, son estudiadas también por Corpas Pastor. En ellas, el autor considera que las más frecuentes son las que llevan el artículo u otro determinante como por ejemplo *la manzana de la discordia*.

2. Locuciones pronominales, que, según el autor, no son muy frecuentes.

3. Locuciones adjetivales, cuyas estructuras más frecuentes son:

> 3.1. Sintagma adjetival. Son estructuras formadas por adjetivo + sintagma preposicional, adjetivo + adverbio o adverbio + adjetivo y adverbio + adjetivo + comparativa. No obstante, el autor insiste en que a veces se pueden unir dos sintagmas y formar un binomio que denomina binomio coordinativo de estructura;
>
> 3.2. Sintagma preposicional. Estas estructuras pronominales sueles llevar la preposición «de» como *de narices*, *de mala muerte*, o *del año de la Tarara*. Estas UF también pueden formar binomios compuestos por preposición + (artículo) + nombre + conjunción + (artículo) + nombre como *de chicha y nabo* o *del pan y del palo*, o también binomios compuestos por de + verbos + conjunción + verbos como *de quita y pon*.

4. Locuciones adverbiales, juntos con las locuciones verbales son, según el autor, las más numerosas. Las más frecuentes son:

> 4.1. Sintagma verbal.
>
> 4.2. sintagma preposicional del tipo preposición + adjetivo y a su vez dividido en subcategorías.

4.3. Sintagma preposicional del tipo preposición + pronombre.

4.4. Sintagma preposicional del tipo preposición + adverbio; y, finalmente, una serie de 4.5) binomios adverbiales.

5. Locuciones prepositivas en las que la unidad lexicalizada con más frecuencia es un sustantivo aunque existen otras variantes.

6. Locuciones conjuntivas, que pertenecen a la categoría de las locuciones nexivas junto a las locuciones preposicionales.

7. Locuciones verbales, que representan la clase más numerosa entre las locuciones. Las estructuras más frecuentes son verbo + complemento, o los binomios coordinativos, aunque existen estructuras particulares como las de polaridad negativa, muy frecuentes según Corpas Pastor (1996) y de polaridad positiva pero con sentido negativo.

8. Locuciones oracionales. Estas representan solo parcialmente una novedad en las taxonomías de las UF. Según el autor, las locuciones oracionales, además de reunir los mismos rasgos fraseológicos generales de las otras clases (fijación, pluriverbalidad, idiomaticidad, institucionalización, etc.), presentan una estructura formal lexicalizada de oración compuesta por un sintagma nominal sujeto y un predicado verbal. Cabe subrayar que las impersonales y un grupo de fórmulas pragmáticas no estarían sujeta a dicha definición. En particular, Marío García-Page distingue entre:

8.1. Locuciones semioracionales, que corresponden a las clausales de Corpas Pastor.

8.2. Locuciones de estructura autónoma, los que Corpas Pastor identifica como enunciados de valor específicos y, finalmente, 8.3) las interjectivas y fórmulas pragmáticas.

Como se ha observado, la clasificación que García-Page aporta en su manual es, hasta cierto punto, novedosa, puesto que ya otros autores habían identificado la locución como un sintagma que forma parte de una oración. Su clasificación es reducida a las locuciones y es excluyente con respecto a las colocaciones y a los refranes. Sin embargo, las aportaciones de este estudioso son interesantes porque profundizan y aclaran algunas de las clasificaciones que expertos precedentes realizan con respecto a las locuciones, única tipología de UF que el autor considera que tienen cabida en el universo fraseológico.

2.3.1.6. Otras propuestas

En esta sección se resumen otras aportaciones a la fraseología española que, a pesar de no ofrecer una clasificación holística del universo fraseológico, son de relevantes para definir el marco teórico de la presente investigación doctoral al haber abordado algunas esferas o características propios de la fraseología española.

La primera autora que se presenta en este apartado sus numerosas aportaciones a los estudios fraseológicos es Carmen Mellado Blanco. La autora ha sido IP de proyectos de investigación estrechamente ligados con la fraseología contrastiva alemán-español[39] y alemán-gallego[40]. Desde el 2007 dirige FRASESPAL (Fraseología contrastiva alemán-español), grupo de investigación interuniversitario que en la actualidad está centrando sus esfuerzos en el estudio de los esquemas combinatorios fraseológicas de estructura [Prep. + Sust.], su valor temporal en alemán y sus correspondencias en español, con especial incidencia en el análisis

[39] Proyecto de investigación: *La estructura idiomática del alemán y el español. Un estudio cognitivo a partir de un corpus onomasiológico* (HUM2007-62198/FILO), Ministerio de Ciencia e Innovación. 2007-2011.

[40] Proyecto de investigación: *Estudio contrastivo sociolingüístico y fraseológico entre el (bajo) alemán y el gallego: su situación en Europa como lenguas menores y análisis de su fraseología como signo de identidad cultural* (PGIDIT04PXIA20402PR), Xunta de Galicia (Promoción Xeral da Investigación). 2004-2006. Proyecto de investigación: *Estudio sociolingüístico contrastivo entre el bajo alemán y el gallego* (PGIDIT02PXIA20404AF), Xunta de Galicia (Promoción Xeral da Investigación), 2002-2003.

de patrones sintagmáticos de estas estructuras en ambos idiomas[41]. El objetivo principal de esta investigación es la publicación de un diccionario bilingüe alemán-español en soporte digital centrado exclusivamente en esquemas fraseológicos.

Su vasta producción se centra principal pero no exclusivamente en la fraseología contrastiva, en aspectos histórico-lingüísticos y dialectológicos de la lengua alemana y española. Su carrera en los estudios fraseológicos empieza con fuerza en 1997 con la defensa de su tesis doctoral titulada *Los somatismos del alemán. Semántica y estructura* en la Universidad de Salamanca. Entre sus principales publicaciones destaca el diccionario bilingüe alemán-español *Idiomatik Deutsch-Spanisch* que en 2013 publica juntos con su grupo de investigación en la editorial Buske (Hamburgo, Alemania). Dicha obra merece ser mencionada por representar sin duda un diccionario pionero en el ámbito de la fraseografía del alemán y español: más de 35.000 entradas en más de 1.200 páginas, siendo la primera obra de tal tamaño en el ámbito contrastivo de estos dos pares de idiomas.

Entre sus artículos más citados destaca *Fraseologismos somáticos del alemán. Un estudio léxico-semántico*, publicado en 2004 por Peter Lang. En su manual, basado en su tesis doctoral, la autora analiza 1385 fraseologismos del alemán que contienen lexemas referidos a somatismos, es decir, partes de la anatomía humana o animal. Mellado describe pormenorizadamente las características composicionales, su idiomaticidad y las variantes estructurales llegando a considerarla como una clase léxico-semántica definida. Como afirma en su introducción, su análisis pretende demostrar que existe una relación directa entre el significado fraseológico y el literal de dichas unidades debido a la escasa opacidad de estas unidaded y estudiar la irregularidad formal y

[41] El proyecto se realiza en colaboración con el Institut für Deutsche Sprache de Mannheim (Alemania), pionero en Lingüística de Corpus en Alemania. para más información véase: <http://frasespal.com/proyecto/resumen-del-proyecto/>.

semántica que caracterizaría a los fraseologismos con el fin de demostrar que dicha irregularidad no tiene que ser considerada como criterio definitorio en sentido absoluto.

La investigadora debate sobre algunas de las principales características de las UF como la idiomaticidad y la opacidad; analiza el significado fraseológico y las relaciones semánticas que se dan en los somatismos centrándose, en particular, en la sinonimia y la antonimia. Asimismo, al estudiar la estabilidad y fijación de las unidades objeto de estudio, llega a afirmar que la capacidad que posee una lengua para producir variantes cambia proporcionalmente en función de lo regular y analítica que sea y evidencia paralelismos en las imágenes metafóricas entre los dos idiomas objeto de estudio. En conclusión, se trata de una obra que presenta una profunda pragmaticidad escrita en un estilo didáctico y divulgativo.

El análisis detallado de otras aportaciones fundamentales a la fraseología española e internacional[42] debe mencionar a Julia Sevilla Muñoz. Fundadora en 1993 y directora de *Paremia*, la primera revista española dedicada al estudio y conservación de los refranes, imparte su docencia en la Facultad de Filología de la Universidad Complutense de Madrid y sus principales líneas de investigación son la paremiología y la fraseología desde una perspectiva contrastiva. Juntos con María Teresa Zurdo Ruiz-Ayúcar, es una de las coordinadoras del *Refranero multilingüe* publicado en la página web del Centro Virtual Cervantes[43], una gran base de datos que recopila una selección de paremias españolas populares, principalmente refranes y frases proverbiales, provenientes de la lengua hablada contemporánea y de textos literarios con su correspondencia en 18 lenguas: alemán, catalán, chino, francés, gallego, griego antiguo,

[42] El 12 de septiembre de 2018, la Associazione Italiana di Fraseologia e Paremiologia (PHRASIS) decide rendirle un homenaje en Roma durante su VI Congreso «Fraseologia e paremiologia: modelli strumenti e prospettive».

[43] Véase: <http://cvc.cervantes.es/LENGUA/refranero/>.

griego moderno, latín, inglés, italiano, portugués, rumano, polaco, húngaro, persa, ruso, árabe y vasco.

Como se apunta en su página web «los objetivos de este refranero son difundir el español a través de los refranes, facilitar la enseñanza/aprendizaje del español como lengua materna y extranjera, proporcionar un instrumento de consulta a los traductores y promover las investigaciones sobre la sabiduría popular». Desde que se empezó en el 2005 hasta el día de hoy, el Refranero se ha convertido en una valiosa herramienta tanto para paremiólogos como para curiosos de los refranes populares porque cada paremia es acompañada por un análisis pormenorizado desde diferentes perspectivas: sinónimos, antónimos, variantes, fuentes, contextos, significado etc. y, por supuesto al ser multilingüe, su correspondencia semántica a otros idiomas o de no existir su traducción literal.

Asimismo, Julia Sevilla es coordinadora del primer Programa de Doctorado en el mundo sobre Fraseología y Paremiología[44] y es directora del Grupo de Investigación Fraseología y paremiología (PAREFRAS). Es autora de una amplia producción científica y ha elaborado una importante taxonomía de las paremias española definiéndolas como «son un enunciado breve, sentencioso, consabido, de forma fija y con características lingüísticas propia» (Sevilla Muñoz, 1993:15). A partir de dichas propiedades, la autora propone su clasificación de las paremias, a saber: 1. Paremias propiamente dichas; 2. Paremias jocosas o irónicas; 3. Paremias científicas; 4. Paremias caballerescas; 5. Paremias publicitarias o propagandísticas; 6. Paremias en desuso, arcaicas o dialectales y de uso restringido; 7. Quasi-paremias; 8. Unidades lingüísticas con algunos rasgos paremiológicos; 9. Unidades no pertenecientes al universo paremiológico que, a veces, pueden participar de algún rasgo

[44] Inicio en el curso 2004-2005, «Estructura y función de las estructuras lingüísticas estables: fraseologismos y paremias», código 299, Universidad Complutense de Madrid).

paremiológico. En su ya clásico artículo, la autora detalla cada una de los grupos, analizando sus características y comparándolas con sus homólogos franceses.

Finalmente, es reseñable su valiosa aplicación didáctica a la enseñanza/aprendizaje del ELE tanto hablado como escrito. Entre sus numerosas aportaciones destacamos el reciente trabajo *Las TIC en la clase de FLE para la enseñanza de refranes y locuciones* (2015) publicado en *Fraseología, didáctica y traducción* (Mogorrón y Navarro Domínguez, 2015). Desde un punto de vista didáctico, la autora plantea la relación y las diferencias que existen entre el refrán y la locución, llegando a proponer diferentes actividades para las clases de francés LE con estudiantes de primer curso de Grado. El objetivo de las actividades es fomentar que el alumnado practique un acercamiento al universo fraseológico ya a partir del principio de su formación universitaria para que aprendan a diferenciar las diferentes unidades fraseológicas.

Otro autor que cabe destacar en este apartado es Antonio Pamies Bertrán, cuya producción científica se centra en particular modo en la fraseología del español, también desde una perspectiva diacrónica, en los aspectos contrastivos de la fraseología española en particular con la francesa, en la prosodia y la fonética en general. Asimismo, es miembro correspondiente de la Academia Norteamericana de la Lengua Española y ha estudiado las variedades del español en las dos Américas. Desde 1998 es miembro del Grupo de Investigación: Lingüística y tipología experimental y tipológica (Cod.: HUM422) de la Universidad de Granada.

Entre sus numerosas publicaciones es de especial interés para el presente trabajo doctoral *Fraseología y metáfora: aspectos tipológicos y cognitivos*, que publica en 2002 con Eva María Iñesta Mena. En el libro, los autores analizan los principios elementales de la tipología léxico-semántica con el fin de elaborar una clasificación para detectar posibles universales conceptuales fraseológicos en 25 idiomas de distintas familias

lingüísticas, sobre todo indoeuropeas[45]. Para ello, parten de un corpus[46] compuesto de 2.000 unidades pertenecientes a diez dominios metas o, empleando la nomenclatura empleada por los autores, modelos icónicos universales por su naturaleza biológica y psicológica, a saber: miedo, ira, injusticia, comer mucho, hambre, delgadez, lejanía, esfuerzo y pobreza.

El objetivo de la investigación es estudiar la relación entre las unidades fraseológicas objeto de estudio y los clichés psicológicos universales señalando los paralelismos que ocurren entre las lenguas. El análisis les permite corroborar que, a partir de un número muy reducido de dominio metas se puede agrupar un gran número de unidades fraseológicas[47]. En conclusión, dicho trabajo ha demostrado que existe una relación entre la universalidad y la especificidad cultural en el universo fraseológico.

Otra investigación interesante y curiosa por su objeto de estudio es el *Diccionario del Insulto* (2000), que Pamies escribe con Juan de Dios Luque Durán y Francisco José Manjón Pozas. Esta obra es fruto de *El arte del insulto* (1997), una colaboración previa de estos tres autores, que examina en clave lexicográfica y retórica el insulto. El diccionario ofrece más de 4000 entradas extraídas de prensa nacional y literatura española y casi todas ofrecen la etimología del lema o de la unidad y ejemplos de uso real. A pesar de que la obra tenga un contenido irónico y/o humorístico, viene a subsanar un déficit en la lexicografía española, ofreciendo un repertorio inédito recopilado de manera estrictamente científica.

En el recorrido que se ha emprendido a través de otros autores y autoras relevantes por sus aportaciones a la lingüística española, se debe dedicar un apartado a M.ª Auxiliadora Castillo Carballo. Dirige, juntos con Juana

[45] Español, francés, portugués, italiano, catalán, rumano, inglés, alemán, holandés, eslovaco, polaco, checo, ruso, ucraniano, esloveno, griego moderno, estonio, húngaro, chino, árabe dialectal marroquí, rifeño, guaraní, ticuna y tanimuca.

[46] «Collection of language data brought together for analysis» (ISO 1087-1, 11:2000).

[47] Cabe destacar que el análisis del corpus lleva los autores a proponer la introducción de otros dominios, a pesar de que no sean universales, a los diez analizado como: objetos, animales, fenómenos atmosféricos, etc.

Santana Marrero (Universidad de Sevilla), el proyecto I+D+I *Patrones Sociolingüísticos del Español de Sevilla* (FFI2015-68171-C5-3-P) y su producción científica se centra en la lexicografía, la lexicología y la fraseología del español, aunque ha trabajado en la enseñanza del ELE. Es miembro del Grupo INTERLÉXICO (HUM 758) ubicado en el Departamento de Filología Española I y Filología Románica de la Facultad de Filosofía y Letras de la Universidad de Málaga cuyas líneas de investigación se centran en la lexicografía.

Entre su producción científica destaca para la presente investigación el artículo *La producción fraseográfica* (2017) en el que, desde un punto de vista lexicográfico, analiza los principales repertorios y compilaciones de obras fraseográficas en un amplio abanico temporal que va desde la Edad Media hasta la actualidad. La autora fija el inicio de dichas recopilaciones en el *Seniloquium* o *Refranes que dizen los viejos,* elaborado alrededor de 1480 y que contiene 497 refranes. El texto es de gran interés académico y gracias al análisis diacrónico se convierte en una fuente valiosa para estudios fraseográficos.

En el artículo *El universo fraseológico. Algunos enfoques* que la autora publica en Revista de lexicografía en 2001, presenta una aproximación a las principales taxonomías de la fraseología española, subrayando similitudes y dificultades. Dado el momento en el que se escribió, la clasificación ha quedado desactualizada pero cabe destacar que se menciona la producción sobre distintos temas de variación lingüística (Castillo Carballo 2018; 2011; 1993).

Otro autor que destaca en este apartado es Pedro Mogorrón Huerta, catedrático y director del Departamento de Traducción e Interpretación de la Universidad de Alicante y especialista en fraseología, análisis contrastivo, principalmente español-francés, y traducción automática. Desde 2005 dirige FRASYTRAM (Fraseología y Traducción Multilingüe), el grupo de investigación adscrito a la Universidad de Alicante con en el

que dedica sus esfuerzos a la elaboración de un diccionario de expresiones fijas multilingüe utilizando el español como lengua de referencia y analizando numerosos idiomas como el francés, el catalán, el italiano, el ruso, el inglés, el polaco y el chino. FRASYTRAM dispone de una interesante página web[48] especializada que ofrece recursos que gran utilidad para estudiar la fraseología.

Desde que defendió su tesis doctoral (1994), ha trabajado numerosos aspectos de la fraseología: desde el tratamiento de las colocaciones en los diccionarios españoles, pasando por el análisis contrastivo de las locuciones verbales en francés y español (2002), el estudio de algunas propiedades como la cristalización (2010a, 2010b), la parasinonimia y la geosinonimia en las expresiones fijas (2015), el orígen paremiológico de algunas unidades (2009) o las variedades de la fraseología en América del Sur (2015).

Entre sus recientes publicaciones destaca, junto con Fernando Navarro Domínguez, el libro *Fraseología, didáctica y traducción* (2015). El manual editado por la serie *Studien zur romanischen Sprachwissenschaft und interkulturellen Kommu-nikation*, dirigida por Gerd Wotjak, presenta en 19 capítulos los resultados de las investigaciones más actuales en fraseología tanto en la didáctica de lenguas extranjeras como en los estudios de traducción e interpretación. Mogorrón es autor de un capítulo en el que analiza diferentes cuestiones relacionadas con la comprensión de las UF, la competencia fraseológica y el tratamiento que se les proporcionan en los diccionarios de referencia en francés y español, haciendo especial hincapié en las variedades diatópicas de los dos idiomas[49].

[48] Disponible en la dirección <https://dti.ua.es/es/frasytram/>.

[49] En particular se hace referencia a las aportaciones fraseodidácticas proporcionadas por Germán Conde Tarrío (Universidad de Santiago de Compostela), M.ª Isabel González Rey (Universidad de Santiago de Compostela), Jean René Ladmiral (I.S.I.T., Universidad de París VII), Inmaculada Penadés (Universidad de Alcalá) y Julia Sevilla Muñoz (Universidad Complutense de Madrid) y de las cuales se hablará más adelante.

Este apartado en el que se han presentado algunas de las principales aportaciones a la fraseología española concluye con Antonio Pablo Zamora Muñoz, autor que ha centrado sus estudios en la fraseología contrastiva italiana española. Zamora es profesor titular del departamento de Traducción e Interpretación de la Universidad de Murcia y director del grupo de investigación *Fraseología, Paremiología y Traducció*n (E0B6-01) de la Universidad de Murcia, grupo fundado en 2009 en el que colaboran nombres de prestigio como David Prieto García-Seco, Julia Sevilla Muñoz, Manuel Sevilla Muñoz, Arianna Alessandro, Carmen Mellado Blanco, etc., que estudia la fraseología y paremiología desde distintas perspectivas: aplicada a la traducción, a la lingüística cognitiva, a la lingüística diacrónica así como a la didáctica[50].

Experto en fraseología contrastiva española-italiana, que investiga desde 1996 tras la defensa de su tesis doctoral *Estudio contrastivo español-italiano de frases hechas,* dirigida por José Antonio Trigueros Cano, y que constituye una de las primeras recopilaciones con enfoque contrastivo italiano-español. Sucesivamente, Zamora se ha centrado más en la zona periférica de la fraseología (Zamora 1999b; 1999c; 2005a; 2006), profundizando en particular en las expresiones de carácter eminentemente pragmático (Zamora 1999a; 2003). En este sentido, el autor comparte la visión de centro y periferia de la que se habló precedentemente (Wotyak, 1988; Ruiz Gurillo, 1998; García-Page, 2008) y elabora una clasificación de las UF en el que el repertorio fraseológico es entendido como un *continuum* no discreto. De manera gradual, parte de un centro prototípico y se amplía hacia unidades periféricas que, a pesar de presentar propiedades específicas que definen una UF comparten, al mismo tiempo, en todo o solo en parte, los rasgos de otras clases.

[50] Para profundizar en las líneas de investigación de FRASEMIA véase <https://curie.um.es/curie/catalogo-ficha.du?seof_codigo=1&perf_codigo=4&cods=E0B6*01>.

Por otra parte, Zamora estudia las unidades fraseológicas pragmáticas (UFP) (2003 y 2005a), que define como «un nuevo tipo del consabido cajón de sastre fraseológico» (Zamora, 2005a: 65) y adquieren un papel práctico en el manual *Hai voluto la bicicletta* (Zamora et al. 2006), que elabora para la enseñanza/aprendizaje de las unidades fraseológicas y de los marcadores discursivos italianos para estudiantes extranjeros. El manual, del que se hablará pormenorizadamente más adelante, representa uno de los pocos ejemplos en los en los que se aborda la fraseología italiana más allá de un mero listado de expresiones y viene a ser una propuesta interesante de acercamiento didáctico a un limitado número de estructuras fraseológicas.

2.4. La fraseología en la lingüística italiana

En la introducción de la serie monográfica n°5 de la Biblioteca Fraseológica y Paremiológica titulada *Fraseología y paremiología: enfoques y aplicaciones*, Vanda Durante resume en pocas líneas los estudios fraseológicos en Italia afirmando que:

> El desarrollo actual de los estudios fraseológicos en Italia, salvo en muy contadas ocasiones, es fruto del empeño de un puñado de *hispanistas* italianos que, en este terreno muy poco abonado, están trabajando y tratando de sortear las innumerables dificultades que plantean la ausencia de una efectiva política lingüística y la escasa voluntad de instaurarla. (Vanda Durante; 2014: 7)

En esta desoladora perspectiva hispanista que la autora presenta, aparecen numerosos claroscuros. En los últimos años, debido al auge que la fraseografía y la fraseología en España y en otros países, también en Italia han nacido congresos internacionales que abordan dichos estudios. A saber:

- *I Jornadas de Fraseología y Paremiología. Perspectivas y aplicaciones en Didáctica, Traducción y Lingüística de Corpus*. Organizado en la

Università Aldo Moro di Bari, coordinado por Vanda Durante, tuvo lugar el 5 y 6 de mayo de 2011 en Bari.

- *Fraseologia e Paremiologia. Seconda Giornata di Studio*. Organizado en la Università Roma Tre por Fernando Martínez de Carnero y Luisa A. Messina Fajardo se desarrolló el 6 y 7 de junio de 2012 en Roma.
- *Fraseologia contrastiva: lingue e culture a confronto*, organizado por la Università di Milano y coordinado por Paola Cotta Ramusino e Fabio Mollica. Tuvo lugar del 9 al 11 de noviembre de 2016 en Milán.
- *Fraseologismi in discorso*, organizado por el Dipartimento di Studi Linguistici e Letterari de la Università di Padova, se realizó del 4 al 6 de junio de 2018 en Pádua.

Además, desde 2014, los fraseólogos italianos disponen de un punto de encuentro anual que, a pesar de su indiscutible juventud, se está rápidamente convirtiendo en un referente en el país. Se trata del congreso anual de la *Associazione Italiana di Fraseologia e Paremiologia* (PHRASIS) que en su breve recorrido ha realizado ya seis congresos internacionales:

- *I Giornata di Studi Internazionali dell'Associazione* PHRASIS. *Paese che vai, usanza che trovi. Riflessioni su fraseología e paremiologia*, coordinada por Carmen Navarro y realizado en Verona el 20 y 21 de febrero de 2014. En las jornadas participaron, entre otros, expertos como Dmitrij Dobrovol'skij, del Institute of Russian Languages of the Russian Academy of Sciences in Moscow; Carmen Mellado Blanco, de la Universidad de Santiago de Compostela; Federica Casadei, de la Università della Tuscia di Viterbo; Antonio Pamies, Universidad de Granada; Cristina Cacciari, Università degli studi di Modena e Reggio Emilia, y Stefania Nuccorini, Università degli Studi Roma Tre.
- *II Giornata di Studi internazionali dell'Associazione* PHRASIS. *Fraseologia e paremiologia: passato, presente e futuro*, que tuvo lugar en Cagliari el 16, 17 y 18 de septiembre de 2015, coordinado por Carmen Navarro.

En comparación con la primera edición, estas jornadas duraron un día más dado el aumento en el número de comunicaciones, que se interpreta como un síntoma del creciente interés para la fraseología en Italia. Por primera vez, los idiomas de las jornadas fueron cuatro: el italiano, el español, el francés, y el inglés.

- *III Congresso Internazionale di fraseologia e paremiologia. Fraseologia, Paremiologia e Lessicografia*, que tuvo lugar en Florencia el 19, 20 y 21 de octubre de 2016. Phrasis se convierte en un congreso internacional y deja de ser una jornada de estudio y por primera vez participa la *Academia della Crusca*, organismo estatal dedito a la normalización de la lengua italiana así como La Universidad de Florencia y el Centro interuniversitario di Geoparemiologia (CIG). Cabe destacar que el nivel de las ponencias ha ido creciendo con respecto a las primeras ediciones. De hecho, las plenarias corrieron a cargo de algunos de los autores que marcaron los estudios sobre la fraseología como Gloria Corpas Pastor, Universidad de Málaga y Alberto Zuluaga, Universidad de Tübingen.
- *IV Congresso Internazionale di fraseologia e paremiologia: Fraseologia e paremiologia, roba da matti!*, tuvo lugar en Bucarest del 27 al 29 de septiembre de 2017. El objetivo de este evento fue reflexionar sobre la fraseología y la paremiología no solo como campos de investigación autónomos, sino como punto de partida para trabajos interdisciplinares e interculturales.
- *V Congresso Internazionale di Fraseologia Internazionale: Lessicalizzazioni "complesse". Ricerche e teoresi*, que tuvo lugar en Catania del 26 al 29 de septiembre de 2018. Con lexicalizaciones "complejas", el evento académico se refiere a las expresiones pluriverbales, expresiones idiomáticas, paremias y expresiones estereotipadas, fórmulas rutinarias, trabalenguas, etc. Se reflexiona acerca de las propiedades semánticas y morfosintácticas de dichas expresiones, del grado de traducibilidad, de su componente

cognitivo, de su representación lexicográfica y acerca de la relación que mantienen con la realidad extralingüística.

- VI Congresso Internazionale di Fraseologia e Paremiologia Phrasis: *Fraseologia e paremiologia: modelli, strumenti e prospettive*, organizado en Roma del 11 al 13 de septiembre de 2019, dedicado a reflexionarr sobre los modelos y herramientas con que se afrontan los estudios fraseológicos y paremiológicos, desde cualquier enfoque científico.
- VII Congresso Internazionale di Fraseologia e Paremiologia Phrasis: *Fraseologia, paremiologia e terminologia*, organizado en Pádua del 3 al 5 de septiembre de 2020. Dicho encuentro está dedicado a la reflexión sobre presencia de las unidades fraseológicas en los lenguajes especializados y su relación con las unidades presentes en los lenguajes generales, entre otros aspectos.
- VIII Congresso Internazionale di Fraseologia e Paremiologia Phrasis: *Fraseologia e paremiologia: modelli e dinamiche*. A raíz del brote de COVID-19, en 2021 el congreso anual de Phrasis fue anulado y pospuesto al 15 y 16 de septiembre de 2022. En ello se debate sobre la modelización de las unidades fraseológicas y paremiológicas, las propiedades semánticas y morfosintácticas en la lengua general y en lenguas especializadas, así como su enseñanza en LE y los problemas traductológicos.

La proliferación de este tipo de eventos científicos coloca a los estudios de fraseología italiana en un momento de auge académico. Como apuntan Autelli y Konecny, «en los últimos años se ha asistido a un *boom* en la investigación fraseológica, incluso en el ámbito italiano. En cuanto a la fraseodidáctica, en otros idiomas existe una serie de contribuciones sobre el tema, mientras que en el italiano se acaba de empezar» (2013: 135).

El presente estudio se adhiere a la concepción amplia de la fraseología descrita en los apartados precedentes. Por ello, antes de analizar el recorrido que la fraseología italiana ha tenido desde sus orígenes, es imprescindible examinar cómo se definen en la lingüística italiana los tres pilares de la fraseología: locuciones, colocaciones y expresiones idiomáticas.

La *Enciclopedia dell'Italiano* (2011)[51], obra de referencia para la lengua italiana y resultado de los esfuerzos del *Istituto dell'Enciclopedia Italiana*[52], dedica entradas detalladas en su enciclopedia online a los lemas «Locuzioni» (locuciones), «collocazioni» (colocaciones) y «modi di dire» o «espressioni idiomatiche» (expresiones idiomáticas). La entrada de las «Locuzioni», firmada por Francesco Bianchi (2010), proporciona la siguiente definición para estas UF:

> Nella terminologia grammaticale tradizionale locuzione è il nome generico che designa qualunque unità linguistica formata da più parole grafiche [...] Nella terminologia grammaticale tradizionale locuzione è il nome generico che designa qualunque unità linguistica formata da più parole grafiche[53]. (*Enciclopedia dell'Italiano*, 2011)

Asimismo, en la entrada se afirma que la característica principal de las locuciones es «la stabilità della sequenza; assai limitata, rispetto ai sintagmi liberi, è la possibilità di cambiare l'ordine dei costituenti di una polirematica». En este sentido, son casi inalterables so pena de una mutación de su significado original. Finalmente, Bianco continúa

[51] Dirigida por Raffaele Simone con la colaboración de Gaetano Berruto, Paolo D'Achille y Giuliana Giusti, a partir del 2102 los contenidos son consultables también online, a través del portal Treccani.it.

[52] Fundado en 1925, el Instituto tiene como finalidad la compilación, actualización, publicación y difusión de la Enciclopedia italiana delle Scienze, Lettere ed Arti». Para más información véase: <http://www.treccani.it/istituto/chi-siamo/>.

[53] «En la terminología gramatical tradicional, locución es el nombre genérico que designa cualquier unidad lingüística formada por más de una palabra gráfica [...] En la terminología gramatical tradicional, locución es el nombre genérico que designa cualquier unidad lingüística formada por más de una palabra gráfica». Traducción propia. Véase: <http://www.treccani.it/enciclopedia/locuzioni_(Enciclopedia-dell'Italiano)/>.

especificando que estas nacen como fenómeno de solidaridad léxica y que se diferencian de las colocaciones porque presentan unidad sintáctica y semántica. Así, las locuciones, a diferencia de las colocaciones, pueden ser consideradas como palabras únicas siendo difícil, o incluso imposible, la inserción de material lingüístico en la secuencia nominal o la sustitución de uno de los componentes léxicos por un sinónimo o hiperónimo.

La entrada «Collocazioni» lleva la firma de Federico Faloppa (2010) y es tan amplia como la precedente. Utiliza la definición de Jezek (2005: 178): «una combinazione di parole soggetta a una restrizione lessicale, per cui la scelta di una specifica parola (il collocato) per esprimere un determinato significato, è condizionata da una seconda parola (la base) alla quale questo significato è riferito»[54]. Según Faloppa, las colocaciones se encuentran en un nivel intermedio entre las rígidas y casi invariables locuciones y las combinaciones libres de palabras. Además, se diferencian de una expresión idiomática porque esta última es, por definición, es una estructura cristalizada cuyo significado se puede deducir por los componentes unitarios que la componen.

El mismo Faloppa (2011) es el autor de la entrada «modi di dire». A pesar de que la entrada tenga esta designación, se especifica que la forma más «técnica» para denominar estos compuestos es «espressioni idiomatiche». Según el autor, en lingüística no se ha aceptado ni elaborado una definición precisa de expresiones idiomáticas. Sin embargo, a continuación afirma que:

> In linguistica si tende oggi convenzionalmente a utilizzare la denominazione espressione idiomatica per indicare semplicemente una locuzione figurata convenzionale, più o meno

[54] «Una combinación de palabras sujetas a una restricción léxica, en la que la elección de una palabra específica (el colocativo) para expresar un determinado significado está condicionada por una segunda palabra (la base) a la que se refiere dicho significado». Traducción propia. Véase <http://www.treccani.it/enciclopedia/collocazioni_(Enciclopedia-dell'Italiano)/>.

> fissa, che può appartenere alla classe dei verbi (tirare le cuoia, vuotare il sacco), dei nomi (patata bollente), degli aggettivi (all'acqua di rose) e degli avverbi (alla bell'e meglio).[55] (Faloppa, 2011)

Se puede concluir que según los autores de las entradas del *Diccionario Treccani* existe cierta equivalencia formal con las definiciones proporcionadas para la lengua española. A continuación, se detallan las principales taxonomías propuestas para las UF italianas con el objetivo no solo de estudiar el marco teórico italiano, sino de averiguar hasta qué punto dicha equivalencia es cierta.

2.4.1. Principales aproximaciones a una clasificación de las UF

Los estudios fraseológicos en Italia no han recibido la atención que merecen y, por ende, no abunda la bibliografía italiana centrada en el tema. Como se verá a lo largo del presente apartado, la perspectiva lexicográfica ha dominado los estudios de lingüística en el país durante las últimas décadas y escasos autores italianos han tratado la fraseología como ciencia independiente. De hecho, cabe destacar que muchos de ellos son hispanistas (Quiroga, 2006; Autelli y Konecny, 2013; Vanda Durante, 2014;).

Por otra parte, aún no se ha impuesto una terminología generalizada (Marello, 1996), lo que ha causado confusión terminológica y una ausencia de consistencia en las definiciones. Por todo ello, Quiroga llega a resumir que uno de los imperativos para la fraseología italiana es «la necesidad de una clasificación aceptada por la mayoría y los estudiosos que proporcione un consenso terminológico» (Quiroga. 2006: 69).

[55] «En la lingüística actual existe una tendencia convencional a utilizar el término espressione idiomatica para indicar simplemente una locución figurativa convencional, más o menos fija, que puede pertenecer a la clase de verbos, sustantivos, adjetivos y adverbios.». Traducción propia. Véase <http://www.treccani.it/enciclopedia/modi-di-dire_(Enciclopedia-dell'Italiano)/>.

2.4.1.1. Simonetta Vietri

Simonetta Vietri publica *Lessico e sintassi delle espressioni idiomatiche. Una tipologia tassonomica in italiano e in inglese* en 1985. Se trata de la primera monografía sobre la idiomaticidad y fijación del italiano y, en ella, la autora pretende demostrar que las UF, que como De Mauro[56] denomina *espressioni idiomatiche* (e.i.) ocupan un papel principal en la lengua desde una perspectiva léxico-sintáctica.

En su pionera monografía, Vietri elabora una clasificación a través de más de 3.000 e.i. recopiladas principalmente a través de diccionarios monolingües italianos. Dicha taxonomía se basa en dos principios: en primer lugar, que las e.i. son analizables como las frases libres; en segundo lugar, que para realizar un examen satisfactorio es necesario una clasificación amplia y detallada. De este modo, Vietri identifica nueve clases de *espressioni idiomatiche*, a saber:

1. C0 + V: *Il piatto piange*;
2. N0+V+C1: *Max perde la tramontana*;
3. N0+V+C1 (a+di) N2: *La paura scioglie la lingua a Max*;
4. N0+V+ Prep+C1: *Max va al fresco;*
5. N0+V+C1+Prep+C2: *Eva porta il broncio a Max*;
6. N0+V+N1+Prep+ C2: *Ugo ha messo Max nel sacco*;
7. N0+V+C1+Prep+C2: *Ida prese la palla al balzo*;
8. N0+V+C1+Prep+C2+Prep+N3: *Ida disse pane al pane a Max*;
9. N0+V+C1+Avv: *Max andrá lontano*.

Como se puede observar, su clasificación se basa exclusivamente en las estructuras transitivas y excluye, por voluntad expresa, todas las estructuras intransitivas así como los proverbios, las fórmulas, los binomios y todas las e.i. con «verbo a suporto» es decir, con un verbo débil

[56] Véase apartado 2.3.1.6. Otras propuestas.

asimilable a un auxiliar como *fare*, essere*, avere* (De Mauro, 1999; Gross, 1981).

En línea con Dardano[57], la autora hace hincapié en que las denominadas frases libres son menos libre de lo que parece al producirse cierta restricción en la distribución léxica, por lo que llega a deducir que estas también conllevan cierto grado de fijación. En cambio, señala que ciertas frases fijas no son tan fijas como se espera y admiten transformaciones de tipo gramatical mientras que, como ocurre también con ciertas frases libres, no aceptan variaciones distribucionales que afecten al sintagma nominal invariable, es decir, no se puede modificar el orden de aparición de sus elementos. Asimismo, observa que ciertas unidades fraseológicas son caracterizadas por la «compresenza di due usi possibili e quindi di un eventuale «scarto d'uso"»[58] (1985: 23). En este sentido, Vietri observa que algunas pueden ser interpretadas tanto en sentido literal como idiomático en función del contexto y de la interpretación que reciben.

En 2014, la autora publica otro interesante ensayo en el que trata de organizar más de 7.000 UF italianas en función de la clase léxico-gramatical que presentan, identificando dos clases específicas: la clase que define C1 se refiere a construcciones formadas por verbo+objeto y las EPC compuestas por construcciones preposicionales con el verbo *essere*. Dicha clasificación es un interesante mapa sintáctica de las UF italianas y dispone de una versión en línea[59] abierta para su consulta. En ella, se presentan dos tablas, a la izquierda categorías de UF estructuradas en función de su construcción composicional básica y, a la derecha, en función de la clase léxico-gramatical que desempeñan.

[57] Véase apartado 2.3.1.6. Otras propuestas.

[58] «Coexistencia de dos posibles usos y, por lo tanto, de una posible "diferencia de uso"». Traducción propia.

[59] Véase: <http://dsc.unisa.it/composti/tavole/sv/table.html>

Estas clases propuestas por Vietri constituyen un valioso recurso lingüístico para quienes deseen investigar en psicolingüística, lingüística computacional y *Natural Language Processing* (NLP), áreas que Vietri está investigando actualmente en la lengua italiana. De hecho, la autora afirma que «pre-constructed lexical resources facilitate idioms retrieval both in the case of "hybrid" and "knowledge-based" approaches to Natural Language Processing»[60] (2014: 137).

Para concluir, cabe destacar que en las aportaciones de Vietri se percibe cierta influencia francesa puesto que la autora se centra en la sintaxis del italiano en el entorno de la metodología del léxico-gramática así como ha sido desarrollada por se ha desarrollado por Maurice Gross en el LADL (Laboratoire d'Automatique Documentaire et Linguistique, CNRS) y de análisis automática de los textos.

2.4.1.2. Federica Casadei

Federica Casadei, Università degli Studi della Tuscia (Viterbo, Italia), en 1996 presenta su monografía *Metafore ed espressioni idiomatiche: uno studio semantico sull'italiano*. Como explica la misma autora en la introducción, el manual se basa en su tesis doctoral (Casadei, 1994) y se justifica por la escasez de estudios sobre la fraseología italiana y, en particular, sobre los aspectos semánticos (1996: 81).

Con el objetivo de utilizar las metáforas conceptuales para describir la semántica de las UF, la autora emplea, como Vietri, el término *espressioni idiomatiche* (e.i. empleando la abreviación propuesta por Casadei) para referirse a unidades fraseológicas con significante fijo «il cui significato non è composizionale, cioè non è funzione dei significati dei componenti»[61] (Casadei, 1995: 28). En este sentido, las i.e. son para la autora locuciones

[60] «Los recursos léxicos preconstruidos facilitan la recuperación de fraseología tanto en el caso de los enfoques "híbridos" como "basados en el conocimiento" del Procesamiento del Lenguaje Natural». Traducción propia.

[61] «Cuyo significado no es composicional, es decir, no está relacionado con los significados de sus componentes». Traducción propia.

fijas con significado convencional no literal. Por otra parte, añade que dichas expresiones, contradiciendo la cristalización que las contradistingue, pueden aceptar cierta variación conservando su productividad semántica (Casadei, 1995: 14-26).

Por todo ello, define un corpus que, desde un punto de vista cuantitativo, sea «rappresentativo dell'insieme delle e.i. [espressioni idiomatiche] italiane» (Casadei, 1996: 88), y lo conforma cruzando repertorios específicos como el corpus de locuciones verbales propuesto por Vietri y diccionarios *generales* como *Il Nuovo Zingarelli* (1988). Dicho corpus ha sido elaborado teniendo en consideración cuatro variables, a saber: naturaleza idiomática de la UF, ser una *espressione idiomatica* verbal, la presencia de sinónimos y/o variantes y, por último, el uso y el registro de la e.i. (1995: 90-96). El resultado es una recopilación de 3.064 expresiones idiomáticas y entre ellas se encuentran tres tipos de e.i.:

1. Icónicas: describen gestos o posturas convencionalmente asociados a cierto significado a través de simbolismo cultural como *toccare ferro; fare le corna; alzare le braccia*).
2. Ejemplos ideales: son e.i. que «descrivono un'azione che rappresenta un esempio ideale del tipo di azione descritto dal significato»[62], y que normalmente representan hacer algo inútil o querer hacer una cosa imposíble como *friggere con l'acqua; lavare la testa dell'asino*.
3. metonimias «puras»: en este caso, no hace distinción entre e.i. metonímicas y metafóricas como *alzare il gomito; scendere in piazza; farsela addosso*.

Por otra parte, la autora defiende que los hablantes no aprenden de memoria cada uso de las expresiones idiomáticas. Al contrario, considera que pueden captar y producir expresiones idiomáticas *ex novo,* al realizar

[62] «Describen una acción que representa un ejemplo ideal del tipo de acción descrita por el significado». Traducción propia.

algún tipo de análisis semántico relacionando el significado literal con el idiomático.

Al igual que Vietri, la autora considera que las expresiones idiomáticas no representan un fenómeno completamente asistemático y de imposible clasificación. A través de un análisis semántico, se pueden reconocer ciertas regularidades y reunir el ingente universo fraseológico en unas clases taxonómicas reducidas[63]. Adoptando el modelo cognitivista[64] plantea campos semánticos determinados para su análisis y propone cuatro campos conceptuales principales en los que agrupar las e.i. (Casadei, 1996: 104):

1. El espacio: en este campo trata las e.i. relacionadas con el dominio espacial y de orientación definiéndolo como «modello dello Spazio».
2. El movimiento y las fuerzas: aquí coloca las e.i. cuyo movimiento es empleado en la metaforización de eventos/entidades de las relaciones entre ellos.
3. El cuerpo: la tercera parte reúne las e.i. que poseen como dominio de origen el cuerpo humano y, en particular, reflejan el estados psíquicos y emotivos.
4. Los dominios culturales: finalmente, al último campo semántico pertenecen las e.i. cuyo dominio de orígen es reconducible a las actividades e interacciones de los seres humanos como el juego, la guerra, el teatro, etc.

La presente tesis doctoral se ha nutrido de dichas consideraciones que, además de haber sido innovadoras en la época, han sido reutilizadas por diversos autores para plantear una didáctica eficaz de las UF en las clases de italiano como lengua extranjera, elaborando el método de manuales

[63] Cabe destacar que Casadei hace hincapié en que, por la propia naturaleza compleja de las e.i., algunas de ellas pueden pertenecer a más de un campo conceptual.

[64] Casadei emplea el método cognitivista «partendo da una situazione mista, come una parola polisemica o un testo, per indagare la rete di metafore concettuali che compaiono e che probabilmente si riferiranno a più domini origene e oggetto» (Casadei, 1996: 101).

en función de los campos semánticos. Con todo, se dejan de lado por primera vez las clasificaciones de impronta gramatical a favor de las relaciones semánticas expresadas por su significado fraseológico.

2.4.1.3. Miriam Voghera

Miriam Voghera plantea la primera clasificación global de las unidades fraseológicas italianas. Colabora con de Mauro en la elaboración del *Lessico di frequenza dell'italiano parlato* (1993) y emplea la terminología *lessema poliremático* o *lessema complesso* (LC) para indicar «sequenze costituite da due o più parole che presentano una coesione interna semantica e/o sintattica tale che possono essere considerate un unico lessema»[65] (De Mauro *et all.,* 1994: 185). Voghera, como ya Vietri y Casadei, los considera como un único lexema compuesto por dos unidades léxicas y nunca superior a un sintagma.

La autora plantea una tabla de cohesión que va desde los LC más parecidos a la categoría de palabra hasta los que se parecen más a un sintagma e individua cuatro grupos (Voghera, 1994: 197-198). A saber:

1. Los que se parecen más a la palabra cuyos constituyentes no se pueden mover ni interrumpir con elementos externos como *alte sfere*.
2. Los LC cuyos rasgos son parecidos a las palabras y cuyos constituyentes no se pueden mover, pero sí interrumpir con otros elementos como *fare finta* (*ho fatto solo finta*).
3. Aquellos LC que poseen rasgos parecidos a un sintagma ya que sus constituyentes pueden moverse, pero no son interrumpibles como *codice civile* (**è quello civile il codice*).
4. Los LC más similares a los sintagmas y cuyos componentes pueden moverse y son interrumpibles por medio de unidades léxicas como *letto a castello*.

[65] «Secuencias que consisten en dos o más palabras que tienen una cohesión interna semántica y/o sintáctica tal que pueden ser consideradas como un solo lexema». Traducción propia.

Voghera identifica que el 64% de los *lessemi complessi* presentes en el LIP no es interrumpible por otro material léxico y el orden de sus componentes no puede sufrir modificaciones sin que se pierda el significado unitario. Por otra parte, el 17% de los LC puede ser interrumpido pero sus componentes poseen un orden fijo; finalmente, solo el 5% de los *lessemi complessi* que conforman el estudio no es interrumpible pero presenta un orden semilibre de sus componentes. En este sentido, la autora reconoce en la lengua italiana una de las propiedades lingüísticas que caracterizan las UF y que ha sido anteriormente descritas, es decir, la fijación fraseológica.

En 2004, publica *Le polirematiche*, capítulo en la monografía *La formazione delle parole in italiano*. Analizando el *Grande dizionario italiano dell'uso* (GRADIT) y el *Lessico di frequenza dell'italiano parlato* (LIP) la autora perfecciona sus teorías precedentes y aborda la primera clasificación holística de las unidades fraseológicas italianas. En este sentido, Voghera clasifica las *polirematiche* en:

1. Nominali: representan el grupo más numeroso, según la autora, que distingue este grupo de LC por su estructura: nombre+adjetivo; nombre+sintagma preposicional; nombre+adjetivo; nombre+nombre. En ellas encuentra que, de las 629 *polirematiche nominali* del LIP, 286 poseen un sustantivo que se repite por lo menos dos veces y que estas son palabras altamente polisemánticas como «*assemblea, camera, colpo, consiglio, carta, diritto, effetto, ente, gruppo, numero, punto, scuola, tempo, stato*» (Voghera, 2004: 64). La polisemia en relación a la frecuencia será una de sus principales líneas de investigación.
2. Pronominali: Voghera individua que las estructuras más frecuentes son constituidas por: pronombre+adjetivo o adjetivo+pronombre; adverbio+pronombre o pronombre+adverbio; pronombre+sintagma verbal; artículo+pronombre; preposición+pronombre; pronombre+conjunción+pronombre.

3. Verbali: las estructuras más frecuentes son: verbo+artículo+nombre; verbo+sintagma preposicional; verbo+adverbio[66]; verbo+adjetivo.
4. Aggettivali: la autora advierte de que no es fácil distinguir estas estructuras de las adverbiales puesto que los adjetivos que componen las LC suelen ser «invariabili e non è possibile ricorrere a marche morfologiche per distinguere gli aggettivi dagli avverbi»[67] (Voghera, 2004: 67). Las estructuras más frecuentes de esta clase son: sintagma preposicional; nombre+conjunción+nombre; adjetivo+adjetivo; nombre+nombre.
5. Avverbiali: las estructuras más frecuentes, a pesar de que puedan presentar variantes, son: preposición+determinate+nombre/adjetivo; preposición+nombre+preposición+nombre/adjetivo; nombre+sintagma preposicional; adverbio+preposición+adverbio; adverbio+adverbio; adverbio+conjunción+adverbio.
6. Preposizionali: las principales son: preposición+nombre+preposición; adverbio+preposición; nombre+preposición; (preposición)+pronombre+preposición. La autora advierte de que en determinadas LC, la preposición puede cambiar dejando inalterado su significado.
7. Congiunzionali: las estructuras más frecuentes son: adverbio+che; conjunción+che, sintagma preposicional+che; frase+che; sintagma preposicional+relativo; (preposición)+pronombre.
8. Interiettive: según la autora, son «formazioni fortemente agglutinate e, di fatto, impermeabili a qualsiasi trasformazione»[68] (Voghera, 2004: 69). Las más frecuentes son: adjetivo+nombre; nombre+adjetivo; nombre+nombre, interjección+sintagma preposicional;

[66] Estas estructuras se corresponden con los *verbi sintagmatici* de Simone (1997) y que, como se verà, no suelen ser considerados como material fraseológico.

[67] «Invariables y no es posible utilizar marcas morfológicas para distinguir los adjetivos de los adverbios». Traducción propia.

[68] «Formaciones fuertemente aglutinadas y, de hecho, impermeables a cualquier transformación». Traducción propia.

preposición+nombre; preposición+nombre+sintagma preposicional; sintagma verbal; verbo+nombre.

La clasificación que Voghera elabora tiene muchos puntos en común con las locuciones que Corpas Pastor (1996) identifica, aunque con algunas diferencias. En primer lugar, la autora no toma en consideración las pronominales que Voghera, y Casares, sí presenta en su taxonomía. En segundo lugar, la autora italiana no analiza las locuciones clausales puesto que superan el sintagma y, por lo tanto, no cumplen con los requisitos básicos de las LC. En tercer y último lugar, si la clasificación de Corpas Pastor estudia desde una perspectiva amplia el universo fraseológico, la de Voghera se limita a una más reducida y hace referencia solo a las locuciones.

Voghera es consciente de la limitación que comporta considerar como únicas variables los aspectos léxico-semánticos o sintácticos, por lo que afirma que, para una clasificación completa, se deben considerar también la semántica, la morfología y la frecuencia de uso. En este sentido, expone que los aspectos semánticos aportan las características principales de los CL siendo la morfología y la frecuencia de uso solo aspectos secundarios.

En función del significado que expresan, es decir, de su semántica, la autora distingue tres tipos de LC (2004: 60-61): 1. con significado pragmático: *grazie mille*, *testa dura*, ecc.; 2. con significado prevalentemente textual o gramatical cuyo significado puede recuperarse exclusivamente por el contexto: *non per niente*, *dal momento che*, etc.; 3. con significado prevalentemente léxico: *ferro da stiro*.

En este sentido, su clasificación es más bien funcional a hacer referencia al significado que las LC expresan que a su función gramatical (verbal,

nominal, ecc.) o, en otras palabras, en su clasificación el aspecto gramatical desaparece dejando paso al semántico.

La autora italiana, como se planteaba con anterioridad, toma en consideración también los aspectos morfosintácticos para elaborar su clasificación, aunque los considere como secundarios. Voghera argumenta que los LC se diferencian entre sí también por su composición interna y por su grado de movilidad, y estos dos factores son los que determinan el fenómeno de la lexicalización de una LC. En este sentido, la académica italiana elabora una hipotética escala de lexicalización en la que las estructuras más propensas a dicho fenómeno son las «estructuras compuestas», seguidas por las «estructuras preposicionales», las «estructuras nominales» y, finalmente las «estructuras verbales».

Por último, retoma la concepción de Dardano (1987) según la cual la frecuencia de uso distingue las *unità lessicali superiori* de las frases sintácticas libres, y considera que para el estudio de las LC es fundamental tener en cuenta dicho aspecto. En este sentido, advierte que la frecuencia no depende de la cantidad de LC que pertenecen a una categoría determinada, es decir, que un número muy bajo de estas estructuras puede ser más frecuentes, como es el caso de los LC *interiettivi*. Asimismo, observa que las LC más frecuentes son las que tienen más tendencia a la polisemia.

En conclusión, se puede afirmar que las aportaciones de Voghera superan y añaden profundización a las de sus predecesores. Las superan porque por primera vez la autora va más allá de aquellas concepciones que solo examinaban a la características léxico-sintácticas de las UF y, por otra parte, une las concepciones de otros autores como Casadei (1995), quienes se ocupan principalmente de los aspectos semánticos de dichas unidades. Sin embargo, no abarca todo el universo fraseológico, ya que quedan excluidas de la clasificación los refranes y las paremias.

2.4.1.4. Raffaele Simone

Raffaele Simone, profesor emérito de la Università degli studi Roma Tre, publica *Fondamenti di linguistica* en 1997, monografía en la que analiza tres diferentes tipos de compuestos: *parole complesse o polirematiche*, *collocazioni* e *idioms*. El lingüista italiano define las *parole complesse* como unidades fraseológicas compuestas por dos o más palabras pero que «funzionano come una parola» (1997: 150). Como se verá más adelante, son las que Dardano define como locuciones verbales y Simone interpreta desde un punto de vista más amplio e incluye en ellas también las *unità lessicali superiori* y los *verbi sintagmatici* de Dardano.

Consciente de que «le voci lessicali non sono tutte monorematiche, cioè formate da una sola "parola" (nel senso convenzionale di questo termine)» (Simone, 2016: 26) elabora una jerarquía propia, tema en el que ha ampliamente centrado sus investigaciones[69]. En este sentido el autor divide las combinaciones de palabras en tres subclases: 1) *combinazioni volatili,* que define como sintagmas compuestos por palabras que se agregan y se desagregan en el discurso cuando este se ha esaurito; 2) *combinazioni preferenziali*, agregaciones de dos o más palabras de varia naturaleza que recurren frecuentemente y definidas por razones estadísticas; 3) *collocazioni,* es decir, secuencias de palabras que suelen presentarse en combinaciones estables y privilegiadas formando sintagmas rígidos o semi-rígidos y que respetan la condición sintàctica formando binomios que siempre incorporan su colocativo. Su condición léxica obliga el binomio a utilizar un colocado determinado o, como mucho, acepta algunos reducidos sinónimos del mismo.

El autor considera que las colocaciones deberían formar una única entrada en los diccionarios, puesto que el vínculo entre colocado y colocativo es tan fuerte que los oyentes pueden deducir la colación con

[69] Véase Simone (2007, 2008, 2014, 2016, 2017).

tan solo escuchar la primera palabra. El problema, como opina Marello (1989), es que a menudo se habla de colocaciones en la lingüística italiana cuando se trata de combinaciones libres puesto que no dan lugar a ningún tipo de unidad compuesta. Por ello la autora prefiere hablar de *collocazioni ristrette* (1989: 62).

Además, Simone introduce otra categoría que define con el término, tal vez demasiado amplio, de *costruzioni*[70]. Dicha categoría atañe a condiciones específicas que se pueden resumir en: 1) *significato costruzionale*: el significado de algunas construcciones es intrínseco a la construcción misma; 2) *marche*: disponen de marcas específicas; 3) *forza pragmática*: algunas proporcionan una información pragmática peculiar; 4) *specificità*: son específicas de un idioma pero, en algunos casos, pueden aparecer similitudes entre dos o más idiomas. Finalmente las *costruzioni* están, a su vez, divididas en *predicative* (con fuerza pragmática y sin fuerza pragmática) y *non-predicative* (en función del nivel gramático-sintáctico)[71].

En recientes e interesantes ensayos Simone (2014; 2016) y Masini (2007) aluden a lo que denomina *ambiente sintagmatico di un elemento*, es decir, la propensión natural de las palabras de formar construcciones privilegiadas con distintos grados de estabilidad y cohesión[72]. Simone considera que estas construcciones «costituiscono per così dire un importante livello ausiliare della competenza lessicale e perciò meritano dizionari che le registrino e le descrivano» (Simone, 2016: 23-24). En cambio, no hay suficientes instrumentos para registrarlas como los

[70] En sentido amplio, dicho concepto es asimilable a las propuestas sobre *Construction Grammar* (CXG) para definir un conjunto de modelos gramaticales en el marco de la lingüística cognitiva y evolución del lenguaje. Para más información véase Goldberg (1995; 2006, 2013), Hoffmann (2013), Corpas Pastor (2015, 2021, 2022), Mellado Blanco (2015), etc.

[71] Para más información véase: Simone (2007).

[72] La autora se inspira al concepto sincleriano de *idiom principle*: «The principle of idiom is that a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analyzable into segments». (Sinclair 1991: 110)

diccionarios de combinaciones de palabras. En este sentido, el autor debatiendo sobre el diccionario del futuro (2016), considera que las *combinazioni preferenziali* y las *costruzioni* pueden, y deberían, representar entradas lexicales, o lemas, autónomos en un diccionario.

Finalmente, Simone emplea el anglicismo *idioms* para denominar enunciados cuyo significado global no se puede deducir por la composicionalidad de los elementos que los componen. Sin embargo, la definición propuesta por el lingüista parece demasiado genérica y no consigue definir con exactitud este tipo de unidades. En este sentido, el autor propone tres subgrupos que, sin embargo, no llega a definir pormenorizadamente: *idioms cristallizzati*, *sintagmi fissi*, *idioms situazionali*.

2.4.1.5. Otras propuestas

En esta sección se analizan las aportaciones italianas de particular interés para el presente marco teórico que se caracterizan por haber realizado aproximaciones parciales o desde un determinado prisma del universo fraseológico. Muchas de estas aportaciones han influenciado los posteriores estudios sobre las UF en italianistica.

Los primeros estudios sobre fraseología italiana provienen de la morfología (Quiroga: 2006), considerada como un procedimiento más para la creación de nuevas palabras. Uno de los primeros lingüistas que se aproximan al análisis de la fraseología con un enfoque pragmático es Maurizio Dardano en su monografía *La formazione delle parole nell'italiano di oggi* (1978). De interés para el presente estudio es la aproximación que el profesor emérito de Historia de la lengua italiana en la Università Roma Tre plantea con respecto a la dicotomía entre las *locuzioni verbali*[73] y los *composti nominali*.

[73] Las que el autor denomina *locuzioni* son, en su mayoría, colocaciones, de acuerdo con la clasificación a la base del presente trabajo (Corpas Pastor, 1996)

El autor afirma que las locuciones «rappresentano un unico costituente semantico» y expresan «globalmente un concetto unico» (Dardano, 2009: 17-18). En su monografía reconoce que el grupo de las locuciones es de difícil clasificación ya que, por ejemplo, algunas pueden subir modificaciones sintácticas como la inserción de artículos, adverbios o adjetivos, mientras que otras no. En este sentido, el profesor reconoce que algunas locuciones son más estables y fijas que otras.

Para el italiano, las locuciones verbales pueden poseer «diversi gradi e varietà di tali locuzioni V+N» (1978: 173) e identifica por lo menos dos tipos de locuciones compuestas por verbo + nombre. El primer grupo estaría compuesto por las locuciones analíticas que corresponden a un verbo simple como *fare l'annuncio*, es decir, *annunciare*, o a equivalentes léxicos como *fare la spesa*. En la actualidad, muchas de las que Dardano considera locuciones verbales de este primer grupo, son denominadas en los estudios de fraseología italiana con el nombre de *costruzioni con verbo a supporto* como *poner por escrito,* es decir, escribir, y que según los exponentes del proyecto LGLI, no pueden ser consideradas como frases idiomáticas.

El segundo grupo está compuesto por aquellas locuciones verbales V+N, cuyos elementos no son unidades léxicas independientes y, al unirse, no existe un verbo equivalente. Un ejemplo podría ser *aver diritto a* ya que el verbo **dirittare* no existe. En este grupo, el profesor incluye las que define *locuzioni stereotipiche*, compuestas por un nombre que representa una unidad léxica autónoma, un arcaísmo o una unidad que, insertada en otra unidad distinta, posee acepciones diferentes.

Con respecto a los compuestos nominales, estudios posteriores los considerarán como *espressioni idiomatiche* y *unità polirematiche*. En particular, Dardano centra su aproximación en los compuestos con base nominal, dividiéndolos a su vez en compuestos nominales endocéntricos y exocéntricos. Al primer grupo pertenecen estructuras formadas por

nombres + adjetivos, que considera muy frecuentes, nombres + di + nombre y, finalmente, nombre + a + nombre (Dardano, 1978: 177-178). Estas unidades, como se ha visto, están presentes en la lingüística española bajo el nombre de colocaciones nominales e incluidas por primera vez por Corpas Pastor como parte integrante del universo fraseológico.

Dardano incluye en los compuestos nominales las unidades formadas por nombre + da + nombre, como *occhiali da sole*, en los que la preposición «da» cumple un papel de complemento circunstancial de finalidad; así como los compuestos nominales constituidos por nombre + nombre. Como se verá más adelante, dicha clasificación será criticada por Luca Serianni, que incluirá este tipo de compuestos entre las *unità lessicali superiori*.

En cuanto al segundo grupo formado por los compuestos exocéntricos, Dardano los diferencia de los endocéntricos al no presentar alguna lexicalización. Sin embargo, Quiroga (2006: 47) apunta que «ofrece(n) algunos ejemplos […] en los que, a nuestro entender, se advierte un significado no composicional y un grado de fijación muy alto». Por otra parte, Dardano ~~el autor italiano~~ define por primera vez el concepto de *unità lessicali superiori* (unidades lexicales superiores), nomenclatura que tomará gran relevancia en la lingüística italiana y que será empleada en futuros estudios de fraseología. Dichas unidades son definidas por Dardano como sintagmas compuestos por dos miembros, un elemento determinado y un determinante. Las principales características que el autor identifica en estas unidades son «La stabilità del rapporto significante-significato, la stabilità della sequenza e la frequenza d'uso»[74] (Dardano, 1978: 170).

[74] «La estabilidad de la relación entre significante y significado, estabilidad de la secuencia y la frecuencia de uso». Traducción propia.

Dardano sostiene que para reconocer dichas unidades es suficiente nominalizar el verbo en el determinado, es decir, emplear el compuesto *«ció che fa che…»* (lo que hace que…) para desvelar si se trata de una *unità lessicale superiore*. Como ejemplo valga el caso de *richiesta di asilo* (solicitud de refugio): lo que hace que el refugio sea solicitado. En ellas, el autor reconoce que, desde un punto de vista formal, las estructuras más frecuentes son las de nombre + preposición + nombre y la de nombre + adjetivo.

En conclusión, las aportaciones de Dardano a los estudios sobre el universo fraseológico italiano fueron aportaciones novedosas para la lingüística italiana y, a pesar de las controversias que se señalan a lo largo de la presente sección, han puesto las bases para futuros estudios fraseológicos.

El segundo autor que se presenta en este recorrido es Tullio De Mauro, profesor emérito de la Sapienza Università di Roma fallecido en enero de 2017. De Mauro es el primer lexicógrafo que analiza en profundidad el tratamiento de la fraseología en los diccionarios monolingües italianos. En 1999, publica su *Grande dizionario italiano dell'uso* (GRADIT) que se propone «rappresentare il lessico della lingua italiana in uso nel novecento tra gli italofoni»[75] (De Mauro, 1999: VII) y, tanto en la introducción como en las notas, el lingüista trata pormenorizadamente las unidades fraseológicas que denomina *espressioni o unità polirematiche*.

Los seis volúmenes del GRADIT facilitan la búsqueda de las *polirematiche* puesto que, a diferencia de otros diccionarios, adopta un criterio único: «i lemmi che sono parole piene (sostantivi, aggettivi, verbi, pronomi, avverbi) e che costituiscono il primo elemento di locuzioni cristallizzate (polirematiche) hanno alla fine della voce una sezione dedicata a tali

[75] «Representar el léxico del idioma italiano en uso en el siglo XX entre los italianófonos». Traducción propia.

locuzioni. Le polirematiche sono lemmatizzate in neretto e in ordine alfabetico»[76] (De Mauro, 1999: XXXIII). En este sentido, el diccionario no solo facilita la búsqueda de estas unidades, sino que ofrece un tratamiento lexicográfico comprensible y sencillo para el utente.

De hecho, de las 131.199 polirematiche que ofrece el GRADIT, hasta 1753 son consideradas como lema autónomo, es decir, representan una única entrada en el diccionario como en el caso de *zig zag, pop artist*, y *bon ton*. Además, al centrarse en la terminología más técnica, se observa que las polirematiche son más frecuentes, factor que el autor justifica puesto que «il ricorso alle polirematiche fa sì che in parecchi ambiti tecnici e specialistici le polirematiche prevalgano sui termini tecnici singoli»[77] (De Mauro, 1999: 1177). Por ejemplo, en el área económica, el GRADIT registra 1.321 polirematiche y 604 términos individuales, o entre las ciencias matemáticas aparecen 1.134 *polirematiche* y tan solo 477 términos individuales. Sin embargo, en las ciencias más jóvenes en las que prevalece la enumeración taxonómica o la libre discursividad (como la filosofía y la historia) el autor afirma que se actúa el proceso contrario, es decir, son más frecuentes los términos individuales que las polirematiche.

Dicha conclusión a la que llega De Mauro es de fundamental importancia para el presente trabajo, puesto que demuestra el papel que la fraseología juega en todo tipo de lenguaje. En este sentido, si el lenguaje especializado predilije la formación de nuevas UF frente al empleo de un solo término, la fraseología sería más productiva en el lenguaje especializado. De ahí la importancia de la fraseología que contribuye a enriquecer el lenguaje en todos los registros y ámbitos.

[76] «Las lexemas que son palabras completas (sustantivos, adjetivos, verbos, pronombres, adverbios) y que constituyen el primer elemento de las locuciones cristalizadas (polirematiche) tienen al final de la entrada una sección dedicada a dichas locuciones. Las polirematiche se lemmatizan en negrita y en orden alfabético». Traducción propia.

[77] «El uso de la polirematica significa que en muchos campos técnicos y especializados, la poilirematica prevalece sobre los términos técnicos individuales». Traducción propia.

En el *Lessico di frequenza dell'italiano parlato* (LIP) elaborado en 1993 por De Mauro et al., el autor define las *polirematiche* y les atribuye tres características principales, a saber: 1. su significado no es descomponible en sus componentes, es decir, la idiomaticidad; 2. no admiten variación léxica o estructurales, es decir, la cristalización; 3. su función gramatical no puede ser deducida por la clase gramatical de las palabras que la componen sino depende del conjunto de sus miembros (De Mauro, 1993: 95).

En el 2000, De Mauro publica *Il dizionario della lingua italiana per il terzo millennio*, cuya fuente es el GRADIT, e incluye más de 30.000 *polirematiche* como lemas independientes. Como en el diccionario, considera solo una muestra reducida de *polirematiche* con lema autónomo puesto que su significado es dado por la suma de sus elementos. En este sentido, en la introducción define las *l*ocuzioni polirematiche como «lemmi costituiti da più parole con un significato unitario da esse non deducibili» (De Mauro, 2000: VII) considerando que la estructura de las polirematiche *es* «analaga a quella delle monorematiche» (De Mauro, 2000: XXII).

En algunos casos, se presenta también la motivación de una polirematica incluyendo así la primera datación y algunas referencias etimológicas. Asimismo, el autor dedica la parte final del diccionario a «strutture formali [...] a uso formale o burocrático», es decir, fórmulas rutinarias. En este sentido, el profesor analiza una serie de «strutture *fortemente* codificate» como las cartas formales, cartas de motivación, solicitud de trabajo, solicitud para concursos, invitación a una boda y modulistica de vario tipo. Tanto la terminología empleada para denominar las UF, como sus postulados serán fuente de inspiración para estudios posteriores sobre fraseología y la elaboración de diccionarios.

Otras aportaciones a la fraseología que se mencionan en este apartado son las de Carla Marello. La autora dedica en su ya clásico manual *Le*

parole dell'italiano (2008 [1996]) unos epígrafes que define *parole formate da più parole*. En este sentido, la autora aborda la fraseología desde una perspectiva lexicográfica, definiendo las UF, que denomina *parole giustapposte* o *sintagmi lessicalizzati*, como

> gruppi di parole graficamente staccate, ma così coesi al loro interno da non poter essere interrotti da altre parole, né cambiati nell'ordine dei componenti e tali dunque da presentarsi alla coscienza dei parlanti come se fossero una parola sola.[78] (Marello, 1996: 2)

La profesora, consciente de que buena parte de los procesos de composición del léxico italiano moderno se explican a través la yuxtaposición de palabras (Marello, 1996: 41), analiza el tratamiento que los *sintagmi lessicalizzati* reciben en los diccionarios monolingüe italianos y los bilingües. En su análisis, Marello reconoce que los diccionarios más modernos reconocen dichas unidades de manera indirecta, es decir, que no hay un *modus operandi* común para definir la fraseología y que, por lo general, no aparecen como lemas sino casi siempre mezclados entre los ejemplos de uso.

Más adelante, la autora italiana centra su estudio en las que define como *collocazioni ristrette*[79], es decir, «combinazione di una o più parole tale per cui, avendo scelto, per esprimere un certo significato complesso, la parola A, la scelta della parola B è lessicalmente vincolata»[80]. Marello

[78] «Grupos de palabras gráficamente separadas, pero tan cohesivas dentro de ellas que no pueden ser interrumpidas por otras palabras, ni cambiadas en el orden de los componentes, y por lo tanto se presentan a la conciencia de los hablantes como si fueran una sola palabra». Traducción propia.

[79] Dicha teoría se corresponde parcialmente con las de Hausman (1989) recuperadas por Corpas Pastor (2021, 2022) para definir las restricciones semánticas de las colocaciones desde la perspectiva de la *Construction Grammar* (CXG): «Collocations can be viewed as partially specified constructions (collocational constructions) that are semantically predictable, contain slots to be filled by a restricted set of lexical items and span various phrasal patterns» (Copras Pastor, 2015: 141).

[80] «Combinación de una o más palabras de tal manera que, habiendo elegido la palabra A para expresar un determinado significado complejo, la elección de la palabra B está limitada léxicamente». Traducción propia. Su definición es inspirada a comunicaciones orales de Mel'čuk en congresos anteriores a la publicación de Clas, Polguère, Mel'čuk (1995).

considera que estas UF pasan desapercibidas por el hablante nativo que las identifica como una normal secuencia de palabras hasta que, al abordar una traducción a otro idioma, se da cuenta de que no se trata de una secuencia libre.

Las collocazioni ristrette son, así, el reflejo evidente de la norma lingüística: a pesar de que el sistema y el uso permita variantes en la colocación, la norma hace que el uso de un determinado colocativo defina o menos una *collocazione ristretta* frente a una combinación libre de palabras y, por lo tanto, convierten a la unidad en un signo de precisión expresiva perfectamente identificable por un hablante nativo.

Desde una perspectiva sintáctica, la autora identifica las principales collocazioni ristrette en cuatro clases, a saber: 1. verbo + nombre; 2. nombre + verbo; 3. nombre + adjetivo; 4. verbo + adverbio. A estos grupos la profesora añade uno más compuesto por verbo + nombre en el que el cambio del artículo o su inserción varían el significado de la unidad.

En un análisis superficial, la definición ofrecida por la autora se podría asimilar a la que la academia española suele ofrecer para las colocaciones. No obstante, esta interpretación no se corresponde con la realidad. De hecho, Marello incluye en las collocazioni ristrette también las que Corpas Pastor (1996) identifica como fórmulas psico-sociales como *Buon Natale*. Por otra parte, según la profesora, las collocazioni ristrette conformadas por un verbo como colocativo predilijen el uso de verbos más específicos y de baja frecuencia.

Finalmente, la autora sugiere adoptar la prueba de la sustitución para identificar una collocazione ristretta frente a una locución idiomática. En este sentido, si la posibilidad de sustituir un elemento sin alterar el significado composicional es nula, se trata de una locución; si es posible pero limitada por la norma y su traducción a otros idiomas no es literal, se

trata de una collocazione ristretta[81]. Por todo ello, las aportaciones de Marello (1996) se han revelado muy significativas para los estudios sucesivos en lexicografía y fraseología, puesto que introduce por primera vez las relaciones semánticas entre los componentes de una UF yendo más allá de la componente morfológica que había predominado hasta entonces.

Por otra parte, para definir con exhaustividad el estado de la cuestión sobre fraseología italiana se debe apuntar la labor de Luca Serianni. En dos de sus gramáticas, Serianni ha tratado, aunque de manera indirecta, aspectos relacionados con la fraseología italiana. El lingüista reconoce tres tipos de unidades: *locuzioni*, *verbi fraseologici* e *unità lessicali superiori*. Entre las locuzioni, Serianni reconoce cuatro subclases dependiendo de la función gramatical que desempeñan:

- las *locuzioni preposizionali* pueden ser de dos tipos: a) compuestas por un adverbio y una preposición (*a proposito; intorno a*); b) por un verbo o un adverbio precedido o seguido por una preposición (*contrariamente a; conformemente a*);
- las *locuzioni congiuntive* son sintagmas complejos que se diferencian de las *locuzioni semplici* aunque su catalogación puede ser complicada (*per il fatto che; dal momento che);*
- las *locuzioni interiettive o esclamative*[82] corresponden a las que Corpas Pastor (1996) identifica con el nombre de fórmulas rutinarias psico-sociales. Están compuestas por grupos de palabras utilizadas en función exclamativa (*Dio mio!*; Santo cielo!);
- las *locuzioni avverbiali* son equivalentes a los adverbios y compuestas por dos o más palabras en un orden fijo (*a poco a poco; tutt'a un tratto*). De acuerdo con la función que desempeñan, pueden ser de varios tipos: de modo, lugar, tiempo, cantidad, juicio, interrogativas,

[81] Véase también el concepto de sustitución paradigmática propuesto por Corpas Pastor (1996).

[82] Serianni utiliza la denominación empleada por Casares (1950), es decir, locuciones interjectivas o exclamativas.

> etc. Una observación interesante del planteamiento de Serianni que merece ser destacada, es que las locuciones adverbiales no siempre pueden ser intercambiables con un adverbio en un mismo contexto. En este sentido, pone como ejemplo la locución adverbial *alla lettera* que corresponde al adverbio simple *letteralmente*. En este sentido, el adverbio en la frase «sono *letteralmente* distrutto» no puede ser intercambiado con su locución adverbial correspondiente «*sono *alla lettera* distrutto».

Con lo que concierne a los verbi fraseologici, Serianni se refiere a los que en la tradición española son denominados como verbos incoativos y que no se suelen considerar como material fraseológico (*cominciare a; iniziare a*). De todas formas, Serianni define un verbo fraseológico como un «verbo che sottolinea una certa modalità dell'azione seguito da un infinito o, in pochi casi, da un gerundio»[83] (Serianni, 1992: 380) y reconoce cinco grupos de verbos fraseológicos, a saber: 1. verbos que indican la proximidad de una acción; 2. verbos que indican el comienzo efectivo de una actividad; 3. verbos que expresan el desarrollo de una acción; 4. verbos que indican la continuidad de una acción; 5. verbos que expresan el final de una acción.

Finalmente, Serianni, como Dardano (1978), habla de *unità lessicali superiori* aunque modifica o amplía el concepto introduciendo los que Dardano define como *composti nominali* como *ferro da stiro* o *guerra lampo*. Son las que Corpas Pastor (1996) articula como colocaciones y que Serianni describe como:

> Parole staccate, ma legate strettamente fra loro nel senso [...] unità lessicali composte da più elementi e non insiemi liberi di parole perché per designare quel particolare oggetto o quel concetto si ricorre sempre alla stessa sequenza di parole, e la successione degli

[83] «verbo que hace hincapié en un determinado modo de acción seguido por un infinito o, en pocos casos, un gerundio». Traducción propia.

> elementi che la compongono non può essere variata, né si può introdurre nella sequenza un qualsiasi elemento.[84] (Serianni, 1992: 676)

La última autora que se presenta en este apartado por sus aportaciones a la fraseología italiana es Arianna Alessandro. La autora es miembro de FRASEMIA, grupo de investigación perteneciente a la Universidad de Murcia y coordinado por Pablo Zamora Muñoz. En su tesis doctoral *Investigación en la acción educativa. Las unidades fraseológicas pragmáticas en la didáctica del español y del italiano como lenguas extranjeras* (2011) centra su estudio en las unidades fraseológicas pragmáticas (UFP), subclase específica de fraseologismos ya identificada por Zamora (2003 y 2005a). Alessandro propone el que denomina como Modelo Fraseológico Pragmático (Mo.Fra.P.), un sistema elaborado con el objetivo de facilitar la didáctica de las unidades fraseológicas pragmáticas en el aula de ELE e ILS.

Se trata de un modelo operativo que permite alcanzar dos objetivos a la vez: por un lado facilitar la enseñanza y el aprendizaje de las unidades fraseológicas pragmáticas a través el empleo de recursos lingüísticos y pedagógicos aplicables en un contexto educativo formal; por otro lado, se trata de una herramienta estructurada que permite favorecer el proceso de investigación-acción en el marco de la didáctica de las UFP en el aula de español como lengua extranjera (ELE) y de italiano como lengua extranjera (ILE). Dado el fuerte enfoque de este sistema hacia la fraseodidáctica, se describe pormenorizadamente en el siguiente capítulo que aborda el estado de la cuestión con respecto a la didáctica de la fraseología.

[84] «Palabras separadas, pero estrechamente vinculadas entre sí en el sentido [...] de unidades léxicas compuestas de varios elementos y no de conjuntos libres de palabras, porque para designar ese objeto particular o ese concepto se utiliza siempre la misma secuencia de palabras, y la secuencia de los elementos que la componen no puede cambiarse, ni puede introducirse ningún elemento en la secuencia.». Traducción propia

CAPÍTULO III

FRASEODIDÁCTICA: ÓRIGENES Y DESARROLLO

3.1. Introducción

Descuidar el componente fraseológico en el proceso de enseñanza y aprendizaje de una LE significa provocar un déficit significativo a la competencia lingüística en general y comunicativa en particular[85], de un hablante (Corpas Pastor, 2003; Alessandro, 2011; Molina García, 2016). La competencia fraseológica va más allá de los aspectos puramente lingüísticos, e incluye el conocimiento pragmático y sociolingüístico de una lengua y, por tanto, es prioritario que se integre en las unidades didácticas de una LE.

Dado el creciente número de estudios sobre didáctica de las lenguas, y que se describirá a lo largo del capítulo, se está consolidando una ciencia independiente que nació como vertiente aplicada de la fraseología: la fraseodidáctica. Su origen, como apunta González Rey (2012) en un artículo en la revista *Paremia* y que marcó un hito fundamental en los sucesivos estudios de fraseodidáctica: «se sitúa en la glotodidáctica o didáctica de las lenguas, en donde emergen las primeras

[85] Desde una perspectiva meramente comunicativa cabría diferenciar entre fraseología significativa dirigida hacia la interacción activa, donde la importancia de la pragmática asume un rol fundamental y, por otra parte, la fraseología estilística, que a menudo puede resultar pedante e innecesaria si es empleada por hablantes no nativos. En este sentido, la locución «Tal persona es *más feliz que un regaliz*» puede ser sustituida con «Tal persona es muy feliz». En cambio, la locución «azul marino» identifica un color específico (HTML: #252440) y no hay otro término para identificarlo, por lo que el hablante deberá disponer de dicha unidad fraseológica en su acervo lingüístico para cumplir con el objetivo comunicativo. fra la fraseologia importante per fini di interazione, per efficacia pragmatica e la fraseologia stilistica che è un di più e che in bocca a un non nativo spesso finisce per apparire fuori luogo e stucchevole.

manifestaciones de interés de docentes de lenguas extranjeras por la parte más viva de la lengua que enseñan» (González Rey, 2012: 67).

A pesar de que, como recuerda Szyndler (2015), el término *phraseodidactik* fue empleado por primera vez por Kühn (1987), habrá que esperar hasta la entrada del nuevo milenio para que se integre en el español, a través de un calco lingüístico, por Larreta Zulategui (2001), en su *Fraseología contrastiva del alemán y el español: Teoría y práctica a partir de un corpus bilingüe de somatismos*. Sin embargo, el autor no profundiza en el término en cuestión, limitándose a señalar un creciente interés para dicho aspecto aplicado de la fraseología.

Unos años más tarde, Ettinger ofrece una definición de fraseodidáctica en 2008 y la identifica como la:

> enseñanza y el aprendizaje sistemático y con base científica de las unidades fraseológicas de las lenguas (Lengua Materna o Lengua Extranjera) con el fin de que se puedan reconocer, aprender y emplear como unidades poliléxicas con significado propio y después aplicar adecuadamente en situaciones comunicativas. (Ettinger, 2008: 96)

Es solo a partir de los estudios de González Rey su grupo de investigación en fraseología y paremiología (FRASEONET) de la Universidad de Santiago de Compostela que se consolida el concepto en España y la fraseodidáctica se empieza a estudiar de manera pormenorizada.

Si por un lado, y como se ha visto en el capítulo II, en los últimos años se ha podido observar una proliferación de los estudios en la fraseología italiana (Autelli y Konecny 2013; Trifone, 2014; Franceschi, 2015; Urzí, 2015, entre otros), por lo que concierne a la fraseodidáctica del italiano, las aportaciones publicadas son todavía muy escasas y, en parte, están ligadas a la tradición más consolidada de la academia de hispanistas. Como apuntan Autelli y Konecny «Negli ultimi anni si è assistito ad un

boom nella ricerca fraseologica, anche in ambito italiano. Per quanto concerne la fraseodidattica, per altre lingue esiste già una serie di contributi sul tema, mentre per l'italiano è ancora agli inizi»[86] (Autelli y Konecny, 2013: 135).

Se observa, por tanto, una sustancial escasez de fuentes primarias académicas que aborden la fraseodidáctica del italiano como lengua extranjera, ausencia debida al hecho de que la fraseología en general, y la fraseodidáctica en particular, son ciencias jóvenes en Italia y, por lo tanto, todavía poco abordadas por la academia en este país. Por otra parte, el hecho de que nos encontremos frente a un tema tan poco estudiado demuestra la originalidad y pertinencia de la presente tesis doctoral y estimula futuras líneas de investigación relacionadas con esta temática.

La fuerte vinculación con la fraseología española es otra de las características de esta área de conocimiento. No es baladí que la primera presidenta de la *Associazione Italiana di Fraseologia e Paremiologia* (PHRASIS) sea Carmen Navarro[87], exponente de relevancia de la fraseología española, con numerosas publicaciones sobre los fundamentos teóricos de la fraseodidáctica del español como lengua extranjera (ELE). Asimismo, la actual presidenta de esta institución es Luisa A. Messina Fajardo, hispanista italo-venezolana y catedrática de la Università degli Studi Roma Tre.

Cabe destacar que, aunque aparezca en esporádicos y recientes trabajos[88], el calco *fraseodidattica* no se ha establecido con fuerza en los estudios de italianistica. Además, cuando aparece, no se hace referencia a la tradición italiana del término, sino que se aborda desde la

[86] «En los últimos años se ha asistido a un boom en la investigación fraseológica, incluso en el ámbito italiano. En cuanto a la fraseodidáctica, en otros idiomas existe una serie de contribuciones sobre el tema, mientras que en el italiano se acaba de empezar». Traducción propia.

[87] Actualmente Luisa Messina Fajardo preside la asociación, con Cosimo De Giovanni como vicepresidente.

[88] Daša Stanič (2014); G. Angela Mura (2015); Termigoni, Susana (2015); Imperiale Ricardo y Elmar Schafroth (2016 y 2019).

tradición española. Por tanto, los estudios sobre fraseodidáctica del ILE son todavía muy escasos y, a pesar de que se registre un creciente interés por dicha rama de estudio de la lengua por parte de los italianistas[89], queda mucho camino por recorrer.

Con el objetivo de definir el estado de la cuestión de la presente tesis doctoral, en este capítulo se establecerá, en primer lugar, el orígen y la definición del término fraseodidáctica. A continuación, se estudiará el telón de fondo de toda investigación, que pretende analizar la didáctica de una lengua extranjera, es decir, el Marco Común Europeo de Referencia para las Lenguas: aprendizaje, enseñanza y evaluación (MCER)[90] elaborado por el Consejo de Europa (2001). En particular, se analizará el MCER desde una perspectiva fraseodidáctica, haciendo especial hincapié en sus fortalezas y debilidades taxonómicas y terminológicas, tomando como referencia las pautas ofrecidas al profesorado y al alumnado en el proceso de enseñanza-aprendizaje que de una lengua extranjera.

Posteriormente, se profundizará en las principales aplicaciones de la fraseodidáctica en las clases de lengua extranjera, y se definirán las diferentes corrientes epistemológicas en relación a la introducción en el aula de lengua extranjera del patrimonio fraseológico de un idioma, definiéndose dos claras corrientes. Por otra parte, se analizarán las principales dificultades que el profesorado y el alumnado encuentran en la enseñanza-aprendizaje de dichas unidades y se identificarán unos patrones comunes de los principales déficits.

Tras esta fase de análisis, fundamentada en un estudio de caso con el fin de profundizar en el objeto de estudio, se tratará de definir las principales cuestiones identificadas por la comunidad académica en su labor de examen de los manuales y métodos generales de ILE y, en particular, se

[89] Véase Zamora et all. (2006), Quiroga (2006), Navarro (2004, 2005, 2008), Alessandro (2011, 2015), etc.
[90] Se ha recurrido a la traducción realizada por el Instituto Cervantes en 2002.

estudiarán dos recientes manuales de italiano como lengua extranjera elaborados con la específica intención de abordar en el aula el universo fraseológico. Por último, se examinará el tratamiento de la fraseodidáctica del italiano y del español en los diccionarios monolingües de italiano y de español, así como en los bilingües italiano-español.

3.2. Orígenes de la fraseodidáctica

A pesar de que «esta faceta de la fraseología ha sido relativamente poco investigada» (Szyndler, 2015: 198), en un artículo publicado en la revista *Paremia* titulado *De la didáctica de la fraseología a la fraseodidáctica,* María Isabel González Rey (2012) define tres momentos claves a través de los cuales poder describir la fraseodidáctica y sus relaciones con la fraseología y la didáctica de una lengua, a saber:

1. *didáctica* de la fraseología[91].
2. didáctica de la *fraseología.*
3. *fraseodidáctica.*

En una primera fase, que la autora denomina *didáctica* de la fraseología, González Rey pone un fuerte acento en la didáctica y, en particular, en el interés demostrado por el alumnado de una lengua extranjera o materna y por el profesorado en el empleo y la didáctica de la fraseología. La segunda fase, denominada didáctica de la *fraseología*, trata el periodo en el que nace el interés de los especialistas, tanto fraseólogos como paremiólogos, hacia el estudio de la fraseología y su aplicación a la enseñanza. Por último, la *fraseodidáctica*, tercera fase del proceso, es la ciencia que estudia la didáctica de la fraseología de forma independiente como una rama de la fraseología aplicada (González Rey, 2012: 67-68).

[91] La cursiva es utilizada por la autora y es empleada para enfatizar sobre una u otra palabra.

Con respecto a la primera fase, la *didáctica* de la fraseología, la autora afirma que se dio a principio del siglo XX por una voluntad propia del profesorado de dar a conocer «la parte más viva de la lengua que enseñan como lengua extranjera» (González Rey, 2012: 68). En este sentido, González Rey enumera en su artículo una muestra reducida, pero significativa de manuales prácticos realizados en dicha época[92].

En cuanto a la segunda fase, la académica realiza un amplio recorrido a través de la didáctica de la *fraseología* española, partiendo de la primera década del novecientos hasta llegar al siglo XXI. Es en este siglo en el que, por primera vez, aparecen trabajos con propuestas distintas de actividades para la enseñanza y el aprendizaje de fraseología. Entre ellos, la autora destaca investigaciones como las conducidas por Bally (1909), Bárdosi (1983) en Hungría, Gréciano (1984), Hausmann (1984), Irujo (1986), Ettinger (1992), Mel´čuk (1993) y Howarth (1998).

Con el inicio del nuevo siglo se publica un gran número de manuales que pretenden abordar estudios sobre la didáctica de la fraseología: Netzlaff (2005), Meunier y Granger (2008), Cavalla (2009). En este contexto, cabe mencionar el relevante papel desarrollado por la Asociación para la Enseñanza del Español como Lengua Extranjera (ASELE) que, desde los años noventa, el Primer Congreso Nacional de ASELE tuvo lugar en Granada en 1989, recopila en sus actas numerosas publicaciones anuales sobre el tema.

La última fase que estudia es *la fraseodidáctica*, término utilizado anteriormente por el germanista Juan Pablo Larreta Zulategui (2001). A diferencia de este autor, González Rey (2012) destaca el valor autónomo de la fraseodidáctica como una rama de la fraseología aplicada. La autora (2012: 76) hace hincapié en que la fraseodidáctica «consiste, ciertamente, en la didáctica de la fraseología de una lengua, pero

[92] En particular, la autora meciona manuales para los siguientes idiomas/pares de idiomas: alemán-francés, francés-rumano, ruso, francés de Canadá, francés de Francia, inglés y, por último, italiano. Para más información véase González Rey (2012: 68).

también en la didáctica de una lengua a través de su fraseología». En este sentido, el fin último de dicha ciencia no es meramente lingüístico, sino semántico y pragmático o, en otros términos, cultural.

Es imposible dominar un idioma sin considerar los aspectos culturales que lo han forjado, puesto que cada idioma es espejo de su cultura y refleja las imágenes mentales de una comunidad de hablantes. Como aludía Serradilla Castaño (2000: 657): «una lengua es reflejo de una cultura y la cultura se puede aprender a través de las expresiones que mejor reflejan la idiosincrasia de un pueblo: las frases hechas […]. Acceder a un idioma sin un conocimiento de estas expresiones se convierte en una tarea imposible». Higueras (1997) apoya esta hipótesis, que atribuye a las UF la función de vehículo de trasmisión del legado cultural.

Por todo ello, en el presente trabajo se utilizará el término fraseodidáctica entendido como una ciencia que atañe a la enseñanza y el aprendizaje de las unidades fraseológicas, tal como son descritas en la clasificación de Corpas Pastor (1996), yendo más allá de lo gramatical, para abordar los aspectos pragmáticos y semánticos de las unidades en cuestión. El fin último de la fraseodidáctica es conseguir que los estudiantes puedan emplear las UF de manera adecuada al contexto, al registro y a la situación comunicativa o, para decirlo con Ettinger, que sepan «en qué situación e con qué intención» usarlas (Ettinger, 2008: 97).

3.3. Marco Común Europeo de las Lenguas

Analizar y reconocer las unidades fraseológicas forma parte del conocimiento más avanzado de una lengua, ya sea materna, segunda o extranjera (Howarth, 1998). Por dicha razón, es de sumo interés para el presente estudio doctoral analizar cómo se aborda la fraseología en la que ha sido la obra didáctica más influyente en las últimas dos décadas: el MCER (2001).

El Marco Común Europeo de Referencia para las Lenguas: Aprendizaje, Enseñanza, Evaluación (MCER) es considerado el estándar europeo para medir el nivel de comprensión, expresión oral y escrita en una determinada lengua. Como expone el Instituto Cervantes en su traducción al español, el MCER pretende:

> […] suscitar una reflexión sobre los objetivos y la metodología de la enseñanza y el aprendizaje de lenguas, así como facilitar la comunicación entre estos profesionales y ofrecer una base común para el desarrollo curricular, la elaboración de programas, exámenes y criterios de evaluación, contribuyendo de este modo a facilitar la movilidad entre los ámbitos educativo y profesional (Instituto Cervantes, 2002: 1).

El MCER establece seis niveles de dominio (del A1 al C2) que permiten comprobar el progreso del alumnado en cada fase de su aprendizaje y a lo largo de la vida. Dichos niveles, a su vez, son divididos en tres grupos:

- Básico (nivel A1 y A2).
- Intermedio (nivel B1 y B2).
- Superior/avanzado (nivel C1 y C2).

Dicho documento es el resultado de más de diez años de investigaciones llevadas a cabo por lingüistas y pedagogos de los países miembros de la Unión Europea y se publica en 2001 por el Consejo de Europa. El MCER desea «proporcionar una base común para la elaboración de programas de lenguas, orientaciones curriculares, exámenes, manuales y materiales de enseñanza en Europa» (2002: 1).

Examinando pormenorizadamente el contenido del MCER, se observa que en varias ocasiones se habla del objeto de la presente tesis doctoral y, para ello, se utiliza una terminología variada: unidades fraseológicas, unidades pluriverbales, frases hechas, expresiones fijas, fórmulas, fórmulas fijas, estructuras fijas, modismos, marcadores lingüísticos de relaciones

sociales, normas de cortesía, expresiones de la sabiduría popular, expresiones que se componen de palabras que habitualmente se utilizan juntas, etc. Ello evidencia la primera dificultad técnica con respecto a la terminología empleada para describir un objeto de estudio.

Dicha pluralidad es sintomática de cierta confusión terminológica que, por un lado, no se ajusta a las investigaciones más recientes en fraseología y, por otra parte, presenta una clasificación oscura y poco clara de los mismos (Sarancho Arnáiz, 2015). Por otra parte, este popurrí terminológico es un reflejo de la pluralidad que caracteriza el universo fraseológico y de las teorías que tratan de ordenar un ámbito tan amplio y complejo como es la fraseología.

Como afirma Sevilla Muñoz, a menudo, «no hay acuerdo en el empleo de una terminología unívoca. Resulta habitual que cada grupo utilice su terminología o que los investigadores traten de innovar creando una terminología» (2012: 22). Dicha pluralidad no solo no aporta nada a los estudios fraseológicos, sino que provoca una cierta confusión a la hora de identificar y caracterizar un determinado segmento oracional.

En el capítulo 5. Las competencias del usuario o alumno del MCER y, en particular, en el subepígrafe *5.2.1.1. La competencia léxica*, se define la competencia léxica como el «conocimiento del vocabulario de una lengua y la capacidad para utilizarlo [que] se compone de elementos léxicos y elementos gramaticales». Entre los «elementos léxicos» el MCER incluye las «expresiones hechas, que se componen de varias palabras que se utilizan y se aprenden como un todo» (Consejo de Europa, 2001: 108). En este sentido, el MCER aborda una primera clasificación (Consejo de Europa, 2001: 108) y ss.) del universo fraseológico que se transcribe fielmente[93] a continuación para posteriormente analizarla en detalle:

• *Fórmulas fijas*, que comprenden:

[93] Las cursivas y la ortotipografía son del texto original (Consejo de Europa, 2001: 108).

– Exponentes directos de funciones comunicativas (véase la sección 5.2.3.2), como, por ejemplo, los saludos: *Encantado de conocerle, buenos días,* etc.

– Refranes, proverbios, etc. (véase la sección 5.2.2.3).

– Arcaísmos residuales; por ejemplo: *Deshacer entuertos, válgame Dios*.

• *Modismos*; a menudo:

– Metáforas lexicalizadas, semánticamente opacas; por ejemplo: *Estiró la pata* (murió). *Se quedó de piedra* (se quedó asombrado). *Estaba en las nubes* (no prestaba atención).

– Intensificadores, ponderativos o epítetos. Su uso es a menudo contextual y estilísticamente restringido; por ejemplo: *Blanco como la nieve* (= «puro»), como opuesto a *blanco como la pared* (= «pálido»).

• *Estructuras fijas*, aprendidas y utilizadas como conjuntos no analizados, en los que se insertan palabras o frases para formar oraciones con sentido; por ejemplo: «*Por favor, ¿sería tan amable de* + infinitivo...?».

• Otras *frases hechas*, como:

– Verbos con régimen preposicional. Por ejemplo: convencerse de, alinearse con, atreverse a.

– Locuciones prepositivas. Por ejemplo: delante de, por medio de.

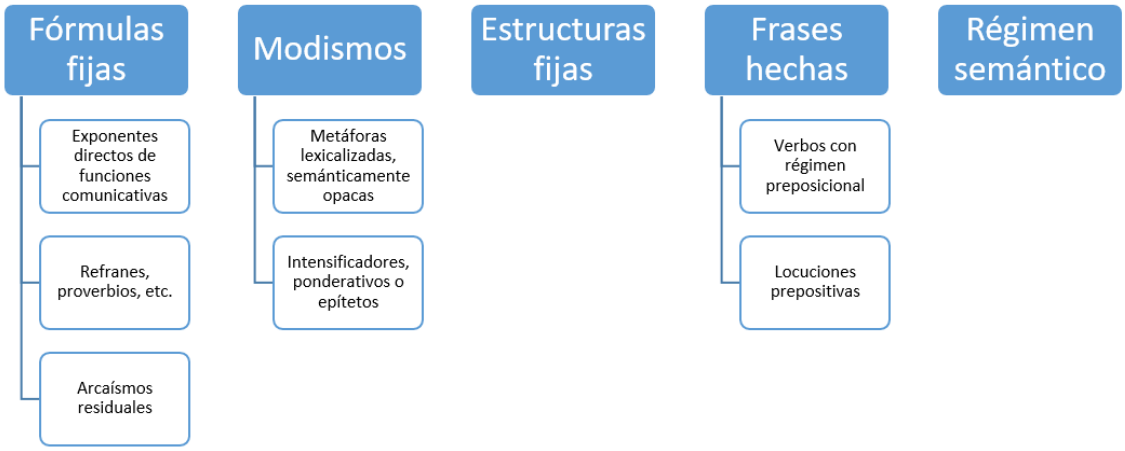
• *Régimen semántico*: expresiones que se componen de palabras que habitualmente se utilizan juntas. Por ejemplo: cometer un crimen/error, ser culpable de (algo malo), disfrutar de (algo bueno).

3.3.1. Sobre el tratamiento de la fraseología en el MCER

La clasificación que aporta el Consejo de Europa en relación al universo fraseológico sufre de cierta inadecuación conceptual y una sustancial ausencia de armonización con respecto a: los criterios ortotipográficos (a veces, se usan puntos para detallar los ejemplos, otras veces comas, otra veces las comillas latinas, etc.); y a las terminología para definir las unidades fraseológicas que, en algunos casos, es demasiada prolija y superficial (véase «expresiones que se componen de palabras que habitualmente se utilizan juntas» (Consejo de Europa, 2001: 108)).

La clasificación propuesta por el MCER divide el universo fraseológico en cinco clases: Fórmulas fijas, Modismos, Estructuras fijas, Frases hechas, Régimen semántico (Consejo de Europa, 2001: 108 y ss.).

Figura 4. MCER Clasificación fraseológica



Fuente: Elaboración propia a partir del Consejo de Europa, 2001.

En primer lugar, se observa que ni la clasificación elaborada, ni la terminología empleada para denominar las cinco clases y las subclases presentan correspondencia formal con ninguna de las aportaciones presentadas en el capítulo II del presente trabajo doctoral. Dicha elección obliga a preguntarse las razones por las que no se han tomado

en consideración anteriores taxonomías, razones que no tienen respuesta en el MCER. Por otra parte, una lectura crítica del MCER, evidencia la presencia de inconsistencias formales tanto en la terminología como en la taxonomía que cambian parcialmente[94].

La clasificación parece mezclar unidades que se caracterizan por estructuras sintácticas y pragmáticas diferentes en una misma clase. Dentro de la clase que denomina Fórmulas fijas se introducen tantos las fórmulas rutinarias (*encantado de conocerle*), como los refranes y las paremias; dentro del grupo de las que denomina genéricamente como frases hechas incluye, sin un criterio claro, los verbos que rigen una preposición y las locuciones prepositivas. Las locuciones, además, están repartidas en distintas clases: las verbales (*estirar la pata*) y adjetivales (*blanco como la nieve*) son consideradas modismos y las adverbiales como frases hechas. En la clase denominada Régimen semántico se incluyen tanto las colocaciones como los verbos frasales (*disfrutar de*), aunque entre estos verbos frasales se encuentren también las frases hechas y, en particular, en la subclase llamada verbos con régimen preposicional (*convencerse de*).

Más allá de la clasificación propuesta, la terminología empleada para definir las clases tampoco parece adecuada. Se emplea el término modismo a pesar de que ya Casares (1950), y luego Corpas Pastor (1996), reconocen que no es un término adecuado y que la lingüística no puede admitirlo, ya que contiene una vaguedad que no permite ceñir sus características. Así, el empleo de términos como estructuras fijas, régimen semántico o frases hechas implica connotaciones semánticas demasiado amplias para identificar una clase de UF: la fijación y la composicionalidad son características intrínsecas de la fraseología y de toda UF.

[94] Véase los apartados: 5.2.2.1. Los marcadores lingüísticos de relaciones sociales; 5.2.2.2. Las normas de cortesía y 5.2.2.3. Las expresiones de sabiduría popular.

Por otra parte, el MCER pone de relieve que el profesorado «tendrá que tomar decisiones concretas respecto a las tareas y las actividades que ha de incluir, qué temas hay que dominar, qué fórmulas, locuciones y elementos léxicos van a exigir que los candidatos reconozcan o recuerden» (Consejo de Europa, 2001: 139). Una vez más se demuestra que se está dejando en las manos de un profesorado que no ha sido formado previamente, como se ha visto en el capítulo anterior, la responsabilidad de la elección de la fraseología, del nivel en el que impartir el contenido y de las actividades que ha de incluir en su programa didáctico. Esta discrecionalidad debería ser evitada mediante su regulación en niveles determinados[95].

Las únicas menciones con respecto a qué unidades impartir en el aula, aunque no se menciona el nivel de referencia asociado, se dan en los apartados *5.2.2.1. Los marcadores lingüísticos de relaciones sociales* y *5.2.2.3. Las expresiones de sabiduría popular* (Consejo de Europa, 2001: 116 y ss.). Por un lado, en el primer apartado se mencionan una serie de fórmulas rutinarias (apertura, transición, cierre, expresivas, etc.):

• Uso y elección del saludo:

- Al llegar a un sitio; por ejemplo: *Hola, buenos días*...
- Presentaciones; por ejemplo: Encantado, ¿cómo estás?...
- Despedidas; por ejemplo: Adiós, hasta luego, hasta mañana…

• Uso y elección de formas de tratamiento:

- Solemne; por ejemplo: Usía, su ilustrísima...
- Formal; por ejemplo: Señor, señora, señorita + apellido, doctor, profesor + apellido.
- Informal; por ejemplo: Juan, Susana, sólo el nombre de pila, sin forma de tratamiento.
- Familiar; por ejemplo: Cariño, amor, mi vida.

[95] Cabe destacar en este sentido el análisis inicial del MCER de Corpas Pastor (2013).

- Perentorio; por ejemplo: sólo el apellido, como: ¡Sánchez, tú, aquí!
- Insulto ritual; por ejemplo: idiota, tonto (a menudo, de forma cariñosa).

• Convenciones para los turnos de palabra.

• Uso y elección de interjecciones y frases interjectivas (por ejemplo, *¡Dios mío!, ¡Venga ya!, ¡Hay que ver!*).

Por otra parte, en la sección *5.2.2.3. Las expresiones de sabiduría popular* se proporcionan algunos refranes, citas, enunciados de valor específico y locuciones. El MCER presenta las siguientes en este orden:

• Refranes; por ejemplo: *No por mucho madrugar amanece más temprano*.

• Modismos; por ejemplo: *A troche y moche. A la pata la llana.*

• Comillas coloquiales; por ejemplo: *To' pa' na'. Me voy pa' casa.*

• Expresiones de:

- Creencias como, por ejemplo, refranes sobre el tiempo atmosférico: *En abril, aguas mil*.
- Actitudes como, por ejemplo, frases estereotipadas del tipo: *De todo hay en la viña del Señor.*
- Valores como, por ejemplo: *Eso no es juego limpio*

Una vez más, el apartado *5.2.2.3.* concluye que cada usuario podrá elegir y evaluar «qué elementos léxicos (expresiones hechas y palabras polisémicas) tendrá que aprender a reconocer y utilizar el alumno, cómo se le capacitará para ello, o qué se le exigirá al respecto; cómo se eligen y se ordenan tales elementos» (Consejo de Europa, 2001: 109). Dicha apertura y ausencia de especificaciones concretas no ayuda ni al docente, ni a los autores de los manuales de lengua extranjera en su labor fraseodidáctico.

Las deficiencias que atañen, no solo a la definición de las unidades fraseológicas, sino también a la asignación de cada UF a un determinado nivel de referencia del MCER, han sido evidenciadas por la comunidad científica de fraseólogos en diferentes artículos y convenios en los que se ha denunciado una falta de rigurosidad metodológica a la hora de abordar las unidades fraseológicas. Entre ellos, destaca Corpas Pastor, quien se hace eco de estas inquietudes y llega a afirmar que:

> «en el MCER se aborda la fraseología de pasada y desde una perspectiva eminentemente ateórica, a tenor de las denominaciones utilizadas, que no son más que simples calcos del inglés, sin mayor precisión terminológica, y que, en muchos casos, inducen a confusión». (Corpas Pastor, 2013: 337)

Es, por ello, que estas debilidades del MCER repercuten inevitablemente en los manuales, las gramáticas y, por lo tanto, en las clases de lengua extranjera (Zuccalà, 2015). Dada la falta de indicaciones por parte del MCER, se ha promovido un debate previamente descrito sobre qué nivel habría que introducir las UF en los manuales y en las clases de LE.

A pesar de que no aparezca explícitamente un análisis pormenorizado de los descriptores[96], el MCER deja patente que la introducción de las UF en la didáctica de las lenguas extranjeras debe ser tratada solo en una fase avanzada del aprendizaje. Por ello, comparte la opinión de Ruiz Gurillo, quien, en el artículo *Algunas consideraciones sobre las estrategias de aprendizaje en la fraseología del español como lengua extranjera* de 1994, firma que «se cree necesario que los manuales de nivel avanzado

[96] Descriptores del Marco de Riqueza de vocabulario: C1: «Tiene un buen dominio de un amplio repertorio léxico que le permite superar con soltura sus deficiencias mediante circunloquios; apenas se le nota que busca expresiones o que utiliza estrategias de evitación. Buen dominio de *expresiones idiomáticas y coloquiales*»; C2: «Tiene un buen dominio de un repertorio léxico muy amplio, que incluye *expresiones idiomáticas y coloquiales*; muestra que es capaz de apreciar los niveles connotativos del significado» (Consejo de Europa, 2001: 109).

de español para extranjeros empiecen a considerar la importancia del caudal fraseológico» (1994: 43).

Por otra parte, cómo se analizó previamente[97], no toda la comunidad científica comparte este punto de vista, y considera que es impensable que un hablante de una LE llegue a un B1 o un B2, niveles en los que el hablante «es capaz de leer artículos e informes relativos a problemas contemporáneos» (MCER, 2001: 31), sin haber sido introducido previamente al estudio de las unidades fraseológicas.

En conclusión, si bien el MCER introduce la enseñanza y el aprendizaje de UF en las lenguas extranjeras a partir de los niveles avanzados, no solo no presenta una clasificación coherente de las mismas, sino que no se emplea la terminología propia empleada por la academia. Todo ello provoca cierta inconsistencia de su aparato teórico. Asimismo, el documento del Consejo de Europa (2011) no aclara el tipo de unidades que se deben introducir en los diferentes niveles dejando mucha libertad, o, mejor dicho, arbitrariedad, al profesorado para que elija de manera autónoma las UF más adecuadas para cada nivel, creando cierta ambigüedad didáctica que merece ser tomada en consideración.

3.4. Aplicación de la fraseodidáctica: «cuándo»

El empleo de la fraseología es un claro indicador del dominio de una lengua (Corpas Pastor, 2013). Desde un punto de vista fraseodidáctico, esta condición es dada por el saber reconocer y utilizar de manera adecuada las unidades; el contexto comunicativo, pragmático y semántico son imprescindibles para su codificación y descodificación y «el mal uso de las UF puede romper convenciones de cortesía, llevar a malentendidos y perjudicar el proceso de comunicación» (Szyndler, 2015: 203).

[97] Véase apartado 3.3. Aplicación de la fraseodidáctica: «cuándo».

Las unidades fraseológicas se rigen en nuestras estrategias mentales para salvar lo que Levinson (1995: 95-98; 2004 [2000]: 29) define como «cuello de botella». En este sentido, la cantidad de información que un hablante pretende proporcionar en un tiempo reducido se ve afectado por el aparato articulatorio del cuerpo humano produciendo un desfase temporal entre el pensamiento y la producción. El sistema cognitivo humano ha elaborado una herramienta para facilitar la necesidad de velocidad, utilizando ciertas estructuras lingüísticas que permiten favorecer la economía del lenguaje. Dicho proceso, en el cual se utilizan moldes semánticos para expresar con un número de palabras reducido un concepto más amplio, no solo es natural en la comunicación entre hablantes nativos, sino necesario.

Desde este punto de vista, para dominar de manera adecuada un idioma extranjero, es fundamental conocer el universo fraseológico de la lengua meta (LM) y saber utilizarlo en el contexto comunicativo más adecuado. Solo si se cumplen estas dos variables se podrá afirmar que el hablante maneja correctamente la competencia fraseológica. A pesar de las dificultades que se encuentran a la hora de enseñar fraseología en una clase de LE, y que se analizarán en el siguiente apartado, el proceso de enseñanza/aprendizaje no puede obviar dichos procesos.

Además, la mayoría de los académicos que trabajan en fraseodidáctica comparten la idea de que la enseñanza de las UF debería de producirse siempre en un contexto determinado. Si las unidades léxicas presentan fenómenos de sinonimia, las unidades fraseológicas presentan características parecidas que atañen al plano pragmático y semántico, y es impensable presentarlas en un contexto determinado, preferiblemente a través de textos auténticos, sean orales o escritos.

Por otra parte, a partir de los ensayos de González Rey se ha ampliamente debatido sobre el «cuándo», es decir, cuál es el momento más adecuado para introducir la fraseología en las clases de una lengua extranjera (LE).

Como sucede con las unidades simples, la comunidad científica coincide en que, para enseñar fraseología, hay que tener en cuenta el conocimiento de los alumnos en relación con el MCER: en cada nivel cabe enseñar/aprender determinadas unidades fraseológicas y no otras.

A la hora de tratar el «cuándo» enseñar las UF, los estudios sobre fraseodidáctica parecen dividirse en dos corrientes dicotómicas. Por un lado, los autores partidarios de que los niveles más adecuados para impartir el universo fraseológico sean el intermedio y el superior y, por otra parte, los que consideran que habría que enseñarlo desde el primer acercamiento al idioma, es decir, desde el nivel más básico.

En una comunicación publicada en las actas el IV Congreso Internacional de la ASELE, Ruiz Gurillo defiende la necesidad de que «los manuales de nivel avanzado de español para extranjeros empiecen a considerar la importancia del caudal fraseológico» (1994: 43). La autora, refiriéndose en particular a las locuciones, considera que el aprendizaje de la fraseología no debe tratarse en profundidad en los niveles elementales, y opina que es aconsejable proporcionar al estudiante una muestra representativa de unidades fraseológicas solo a partir de los niveles avanzados (Ruiz Gurillo, 1994; 2000).

En su monografía dedicada a las locuciones del español en 2001, reserva un apartado a la enseñanza de este tipo de UF y vuelve al tema del «cuándo», afirmando que, a pesar de que los métodos comunicativos de la enseñanza de las LE tiendan a proporcionar el universo fraseológico desde el nivel elemental, «las paremias y las locuciones suelen dejarse para más adelante» (2001: 89). En este sentido, la autora propone que las fórmulas rutinarias se inserten en los niveles más básicos, mientras que las locuciones y las paremias en los niveles intermedio y avanzado.

Comparten la opinión de la autora otros académicos como Koike (2001), quien considera imprescindible que los estudiantes de los niveles intermedios aprendan las colocaciones porque plantean problemas de

codificación y descodificación. López-Jiménez (2013: 338) también comparte esta postura e, incluso, la supera, al considerar que lo mejor sería empezar a ofrecer el entero universo fraseológico a partir de los niveles intermedios.

Por otra parte, existe una serie de estudios que defienden la introducción de las UF desde los niveles más básicos. María Isabel González Rey (2006) explica que, desde un punto de vista fraseodidáctico, introducir la fraseología a partir de los niveles superiores no es la mejor elección. Según la autora, lo más adecuado es introducir las UF gradualmente desde los primeros niveles del MCER, ya que es impensable que un hablante de LE llegue a un B1 o un B2, niveles en los que el hablante «es capaz de leer artículos e informes relativos a problemas contemporáneos» (MCER, 2001: 31), sin haber sido introducido previamente al estudio de las unidades fraseológicas.

De la misma opinión son autores como Marello que, hablando de las colocaciones considera que «le collocazioni ristrette vanno apprese fin dal primo anno di studio di una lingua straniera»[98] (Marello, 1989: 63). También Quiroga considera que ya a partir del nivel básico se deben introducir las UF, en particular aquellas «más frecuentes y más sencillas por sus características formales, semánticas y pragmática y dejar para los niveles superiores aquellas que presentan más dificultades» (Quiroga, 2006: 184). De acuerdo con este argumento, Szyndler plantea que las UF, «a pesar de su complejidad interna y de su heterogeneidad conceptual, se pueden introducir en el aula de manera productiva y eficaz, incluso a partir del nivel inicial» (Szyndler, 2015: 199).

Serradilla Castaño observa cómo el alumnado de ELE, incluido el que acaba de relacionarse al idioma español, se ve naturalmente fascinado

[98] «Las colocaciones restringidas deben aprenderse desde el primer año de estudio de un idioma extranjero». Traducción propia.

por ciertas locuciones frente a otros aspectos gramaticales del idioma. En este sentido afirma que:

> «es curioso observar cómo los estudiantes que llegan a España para aprender nuestra lengua aprenden normalmente antes "estar de puta madre" que la conjugación completa del presente de indicativo de SER y ESTAR; así que si son capaces de aprender expresiones como ésta ¿por qué no darles alguna otra frase que pueda serles también útil?». (Serradilla Castaño, 2000: 663)

El presente trabajo comparte esta última posición teórica, según la cual las unidades fraseológicas deben de enseñarse/aprenderse desde el primer momento de estudio de una lengua extranjera y sería propicio introducirlas en el aula de forma paulatina. El alumnado podría de este modo familiarizarse con dichas unidades ya desde el principio de su formación, porque son necesarias e imprescindibles para la comunicación con nativos.

De hecho, en el presente estudio doctoral se hace propia la idea que el profesorado puede y debe aprovechar las propiedades lingüísticas de las unidades fraseológicas[99] puesto que su características formales pueden ser de ayuda para la labor fraseodidáctica: algunas son tan frecuentes que es obligatorio que el alumnado las reconozca desde el principio (*buona serata*), otras menos frecuentes pero muy adecuadas para determinadas situaciones que se establecen dentro y fuera del aula (*in bocca al lupo*); otras son tan idiomáticas que estimulan el alumnado (*avere le mani bucate*), otras, en cambio, son tan poco idiomáticas que el estudiante puede llegar a decodificarla de manera autónoma (*occhiali da sole*); en otras la motivación de la UF es exclusivamente propia de la cultura italiana y, por ello, habrá que enseñarlas y aprovechar el factor motivacional (*dare pan per focaccia*), otras son

[99] Véase Capítulo II, apartado 2.1.2. Propiedades lingüísticas.

europeísmos fraseológicos comunes a varias culturas y, por lo tanto, interpretadas por el estudiante como un calco de su lengua materna y fácilmente asimilables (*tallone di achille*).

Por todo ello, si bien es cierto que «el requisito de haber alcanzado un nivel intermedio o superior es sin duda ineludible para captar ciertos matices semánticos y utilizar las locuciones en contextos lingüísticos adecuados» (Gutiérrez Quintana, 2004: 24), no se puede preservar al alumnado de aprender dichas unidades fraseológicas, ya desde los primeros contactos con la lengua estudiada.

Tras establecer el «cuándo» enseñar las UF, se considera que aún más importante e imprescindible es averiguar «cuáles» son las unidades fraseológicas que se deben enseñar en cada nivel y qué criterios se deben atender para discernir, objetivo principal del presente trabajo doctoral. Sin embargo, la ausencia de estudios científicos que indiquen al profesorado qué UF enseñar en cada nivel del MCER y cómo presentarlas al alumnado (Penadés, 1999: 33) plantea lo expuesto en el plano teórico sin materialización en las clases de LE.

3.5. Dificultades de la fraseodidáctica en las clases de LE

Es necesario relacionarse con un idioma, tanto el materno como un extranjero, como algo vivo y cambiante y, por ello, es deber del profesorado reconocer dicha característica de la lengua, registrar los cambios y transmitirlos al alumnado. Marello afirma que «come gli esami, la formazione dell'insegnante di lingue non finisce mai. Quando si crede di aver descritto e compiuto, per esteso, sistematicamente, un percorso, la lingua da insegnare è già cambiata perché viva»[100] (2006: 10).

El léxico es, con toda probabilidad, la parte de la lengua más sujeta a modificaciones en cuanto se adapta a la sociedad y a sus necesidades

[100] «Como los exámenes, la formación del profesorado de idiomas nunca termina. Cuando crees que has descrito y completado, de forma completa y sistemática, un camino, el lenguaje a enseñar ya ha cambiado porque viva». Traducción propia.

reflejando el medio de los hablantes. En este sentido, la International Standard Organization (ISO 704, 2000: 31-35) individua tres patrones para la creación de neologismos, es decir, la formación de nuevos términos: creación de nuevas formas, creación a partir de formas ya existentes; y préstamos/calcos a partir de otros idiomas[101]. Teniendo en cuenta estos tres orígenes, la fraseología respeta dichos patrones siendo una ciencia que guarda estrechas relaciones con la lexicografía y la terminología.

Esta naturaleza líquida, en un sentido *baumiano* del término, la convierte en una ciencia complicada de enseñar y aprender. De hecho, se puede afirmar con Olímpio que «hay […] un consenso al respecto: todos los investigadores opinan que las UF son difíciles de enseñar y que a cualquier aprendiz de una LE le cuesta dominarlas» (2006: 6). A pesar de las dificultades que el alumnado puede encontrar a la hora de aprender fraseología, a lo largo de este apartado se verá que el profesorado de lengua extranjera tiende a no enseñarla de manera estructurada en sus clases y se profundizará en las razones que llevan a dicha limitación.

En un capítulo de *La Enseñanza del Español como Lengua Extranjera: del Pasado al Futuro*, Forment Fernández (1998) detalla cuáles son, en su opinión, las principales razones por las que la fraseología se trata solo parcialmente en las clases de LE. La autora evidencia la dificultad de comprensión reconducible a las peculiaridades morfosintácticas de las UF, que suponen un obstáculo para el alumnado. Por otra parte, también el profesorado es, a menudo, incapaz de explicarlas y/o clasificarlas, al no disponer de materiales didácticos para su enseñanza. Por estas razones, se observa que la dificultad es bidireccional: por un lado, atañe al alumnado y, por otro, al profesorado.

En este binomio profesorado-alumnado se identifican una serie de problemas que merece la pena analizar de manera individual a pesar de

[101] Para más información véase la ISO 704: (2000): *Terminology work. Principle and methods*, Geneva, International Standard Organization.

que sean parte de un todo. Desde el punto de vista del alumnado de una LE, la primera controversia que se le plantea es la descodificación: reconocer en un texto oral o escrito que el segmento oracional en cuestión no representa una asociación libre de palabras, sino una unidad fraseológica de alguna manera institucionalizada en el idioma meta. Tanto el estudiante como el traductor[102], que traduce a su lengua materna, muchas veces no sabe identificarlas, y ello puede causar problemas, que atañen directamente al nivel semántico, causando un error, y pragmático, interrumpiendo la comunicación.

Más allá de la descodificación de las UF, la segunda dificultad que se presenta a los estudiantes de LE es la producción. Una vez que se ha comprendido el significado de una UF, es imprescindible utilizarla correctamente, es decir, en un contexto y registro adecuado. Para ello, es necesario que el alumnado se familiarice con ella, juegue con la unidad y la emplee en contextos reales.

Con el objetivo de favorecer el proceso de aprendizaje en las clases de LE, es habitual recurrir a equivalentes formales de la LO para conseguir la reproducción de la misma imagen mental que en la lengua materna del aprendiente. No obstante, dicho proceso de equivalencia explicado pormenorizadamente por Corpas Pastor (2003)[103], no siempre es tan sencillo como podría parecer a primera vista, y a menudo esconde problemáticas significativas. Su similitud conceptual no es más que aparente en algunas ocasiones: falsos amigos fraseológicos, equivalencia solo parcial en el plano morfológico, pero no en el significado o viceversa, etc. Por otra parte, el alumnado no puede limitarse a parafrasear una unidad fraseológica, puesto que, en muchos

[102] Numerosos son los estudios sobre fraseología y traducción en los que se evidencian los peligros y problemáticas que se encierran en esta disciplina, así como las posibles técnicas de traducción de las UF. Entre ellos recordamos Sevilla Muñoz (1997); Recio Ariza (2011; 2013); Szyndler, Agnieszka (2014); Fernández Sola et all. (2015), Corpas Pastor (2001, 2015b, 2016, 2018, 2020); etc.

[103] Para profundizar en el concepto de equivalencia véase Corpas Pastor (2013: 217-218).

casos, poseen un valor semántico intrínseco cuya paráfrasis limita o anula por completo.

Si el alumnado es el destinatario último de la fraseodidáctica, el profesorado es el medio a través del cual el proceso se realiza y su labor tampoco es sencilla y libre de escollos. En este sentido, suele pasar que «el profesor que decide trabajar en el aula las UF se encuentra desorientado respecto a las combinaciones que tiene que trabajar y de acuerdo con qué criterios» (Carmen Navarro, 2004: 101). Esta confusión ocurre por varias razones y una de las causas es, sin duda, la escasez esencial de investigaciones que apoyen al profesorado en su labor de enseñanza de la fraseología.

En un estudio de caso realizado por Zuccalà (2018), fruto de una estancia de investigación realizada en la Pontificia Universidad Católica de Ecuador[104], se examinó el tratamiento que la fraseología recibe en las clases de italiano como lengua extranjera (ILE) en dos ciudades: Málaga (España) y Quito (Ecuador). Para ello, se proporcionó un cuestionario *ad hoc* al profesorado de ILE perteneciente a distintas instituciones de enseñanza de las dos urbes con el objetivo de obtener información cuantitativa y cualitativa respecto a la utilización de la fraseología italiana en las clases de italiano como lengua extranjera. El panel de expertos que participó en esta investigación estaba compuesto por veinte profesores y profesoras de ILE, y el objetivo era analizar la información obtenida en cada ciudad para posteriormente compararlas entre ellas e identificar patrones comunes sobre el tratamiento de la fraseología en las clases de ILE en las dos urbes objeto de estudio.

El análisis de los resultados de los cuestionarios proporcionado al profesorado de ILE en ambos países, demostró que, en primer lugar, los

[104] Zuccalà Daniele (2018), "Fraseodidattica e Italiano LS. Studio di caso a Málaga (Spagna) e Quito (Ecuador)", Coord. Benucci, Elisabetta (2018), *Fraseologia, Paremiologia e lessicografia III Covegno di PHRASIS*, Roma:
ARACNE (Topoi 5).

expertos que componían la muestra consideraban imprescindible introducir la fraseología en las clases de ILE a partir de los niveles más básicos del MCER (A2), participando de todo el proceso de aprendizaje desde el inicio.

Por otra parte, de acuerdo con Saracho Arnáiz, «el profesor debe tener una noción general de las características externas e internas de las UF, ya que es fundamental a la hora de seleccionar las unidades para un grupo/nivel de alumnos determinado y para elaborar los materiales de enseñanza-aprendizaje» (Saracho Arnáiz, 2015: 923). Sin embargo, el estudio demostró que el profesorado de ILE no había recibido formación en fraseología y, por ende, desconocía materiales específicos para impartir fraseología en el aula. Todo esto provocaba que no planificaran los contenidos fraseológicos en las unidades didácticas y que se guiasen por dos aspectos que escapaban de su control: por un lado la estructura impuesta por los manuales didácticos de ILE y, por otro lado, la conversación natural con el alumnado.

Teniendo en cuenta que los informantes opinaron que los manuales de ILE eran inadecuados para impartir fraseología, fue llamativo que el 60% de ellos afirmase que se guiaba por estos recursos didácticos para enseñar contenidos fraseológicos a su alumnado, mientras que el 30% del profesorado se regía por la conversación natural que se articulaba en el con los estudiantes.

Dichas aseveraciones, demostradas de manera empírica a través de un estudio de campo, corroboran la opinión de Saracho Arnáiz, quien afirma que «los profesores de ELE no tienen una formación teórica suficiente para reconocer los límites del ámbito fraseológico y, como consecuencia, no le pueden otorgar un lugar de importancia en el aula» (2015: 921). Asimismo, la autora, de acuerdo con los resultados de la investigación (Zuccalà, 2018), reconoce que «en ocasiones, los contenidos fraseológicos aparecen en la clase, pero de forma accidental, sin que

haya habido un plan y un deseo específico de que una u otra unidad esté presente para cumplir una finalidad de enseñanza-aprendizaje» (*idem*). Se puede concluir que el diseño de la didáctica de la fraseología por parte del profesorado en el aula de italiano como lengua extranjera se rige por una sistemática arbitrariedad y discrecionalidad.

Por todo ello, de acuerdo con Francisco Núñez-Román (2015: 164), «es necesaria la formación del profesorado en este campo, pues se trata de un ámbito de la lingüística de reciente desarrollo y aún poco conocido en la formación del profesorado». En primer lugar, es necesario proporcionar formación sistemática en fraseología al profesorado de LE con el objetivo de que pueda reconocer y poner en valor estas estructuras en el aprendizaje de un idioma; y, en segundo lugar, es necesario elaborar material fraseodidáctico específico por idioma y nivel que permita favorecer el desarrollo de la competencia fraseológica de los aprendientes.

Enseñar la fraseología «no es *coser y cantar* para el docente, sino un proceso lleno de dificultades de diversa índole» (Szyndler, 2015: 212). Con todo, no es imposible y en las últimas décadas se ha asistido a la creación de diferentes tipos de manuales que ponen la fraseología en el centro del aprendizaje. A continuación, se describen los principales estudios realizados sobre manuales de ILE con un prisma fraseodidáctico.

3.6. La fraseodidáctica en manuales de ILE y ELE

Los manuales de enseñanza de lenguas extranjeras desarrollan un papel fundamental tanto para el profesorado, dado que representan una guía para desarrollar su labor en línea con el MCER; como para el alumnado, ya que le proporciona un contacto directo con la lengua meta. Sin embargo, y como afirma la Comisión Europea en su comunicación para el Consejo y el Parlamento Europeo, «los sistemas de formación del

profesorado no siempre están bien preparados para satisfacer estas nuevas demandas» (2007)[105].

Los estudios publicados sobre el empleo de la fraseología en los manuales de enseñanza de español como lengua extranjera (ELE) han demostrado que, si por una parte el papel de la fraseología ha adquirido más relevancia en los últimos años, sigue habiendo todavía muchas lagunas, en particular, en los niveles básicos y en relación a la elección de las UF a introducir en los libros (Daša Stanič, 2014). Por lo general, los estudios que han analizado el tratamiento de las UF en los manuales de lengua extranjera han evidenciado que el espacio que se le dedica a dichas unidades es en su mayoría insuficiente (Ruiz Gurillo, 2000).

Al tratar de describir la atención que los manuales de ELE ofrecen a las unidades fraseológicas, se observan problemas similares a los que se han descrito para el italiano. Gómez Gonzalez y Ureña Tormo destacan que:

> No resulta infrecuente encontrarlas dispersas en el manual sin una organización metodológica o al final de una unidad por cuestiones de afinidad temática. También se incluyen a menudo como un dato anecdótico a propósito de un texto o una actividad, o como parte de alguna lección dedicada al español coloquial [...]. En cualquier caso, en ninguna de las circunstancias indicadas, la enseñanza de estas unidades parece quedar integrada en los contenidos del curso ni trabajarse según un planteamiento fundamentado. (Gómez Gonzalez y Ureña Tormo, 2014: 315)

López-Jiménez (2013: 344) ha individuado tres problemas principales en los manuales en relación al tratamiento del universo fraseológico. En primer lugar, los criterios utilizados en la elección de las UF son, a menudo, aleatorios; en segundo lugar, la tradición de la enseñanza que considera los niveles más altos como los adecuados para impartir UF; y, por último,

[105] Para más información véase: «Mejorar la calidad de la formación del profesorado» disponible en el portal de EUR-Lex <http://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex%3A52007DC0392>.

la ausencia de actividades comunicativas desde la fraseología que caracterizan la mayor parte de los ejercicios.

Con respecto al ámbito de la fraseodidáctica italiana, el tratamiento de las UF en los manuales de italiano como lengua extranjera (ILE) es aún más deficitario que en el caso español. Como registra Quiroga «las investigaciones en torno a la enseñanza de las UUFF en el aula de italiano como LE son muy escasas» (2006: 184). Desde una perspectiva más teórica, la misma autora intenta subsanar ese déficit estudiando el tratamiento que reciben en nueve manuales de italiano como LE en los tres niveles del aprendizaje: dos manuales para el nivel elemental, cuatro para el intermedio y tres para el avanzado. Quiroga se propone analizar «cómo aparecen distribuidas las UUFF y cómo se tratan, qué denominaciones se les da y si se hace alguna referencia específica en el prólogo» (2006: 187). Finalmente, la autora ofrece «una valoración de cada nivel» (*idem*).

Las conclusiones de su estudio son particularmente interesantes a la hora de averiguar la importancia que se da a las UF en cada nivel. Con respecto al nivel básico o elemental, la autora observa que en los dos manuales analizados «apenas se da importancia a las UF (2006: 188), y que la única alusión que se les hace concierne a las fórmulas rutinarias más básica. En particular, la autora se refiere a las fórmulas rituales de saludo y despedida y algunas fórmulas expresivas.

El análisis de los cuatro manuales para el nivel intermedio ofrece resultados igualmente interesantes. Además de identificar cierta confusión terminológica al denominar las UF, los manuales priorizan las locuciones verbales frente a otras locuciones u otras esferas de la fraseología. Además, no solo no suelen ofrecer aclaraciones semánticas y pragmáticas, sino que no consideran los contextos de uso. Por otra parte, observa que no se suele dedicar un apartado específico en la elaboración de la unidad didáctica a dichas unidades y que no se

introducen de manera sistemática. En conclusión, la académica afirma que, a pesar de que se tenga en consideración el universo fraseológico, se hace de manera desestructurada.

Con respecto al nivel superior, la autora observa que el tratamiento que los manuales reservan a la fraseología es «ciertamente superficial y a veces confuso» (2006: 198), y que incluso algunos manuales ofrecen menos UF que en los niveles intermedios. Además, resalta la ausencia total de ejercicios relacionados con las colocaciones.

Otro trabajo sobre análisis de manuales de ILS es el que realiza Daša Stanič (2014). La autora analiza trece recientes manuales de ILE empleados en el sistema educativo esloveno. El análisis de los manuales deja patente una serie de observaciones relevantes para el estudio de la fraseodidáctica italiana y que se exponen a continuación. En primer lugar, averigua que «le UF non vengono incluse nei libri di testo in modo uniforme» (Stanič, 2014: 81), es decir, que existen diferencias entre los métodos tanto en un mismo nivel como en los distintos niveles establecidos por el MCER y las principales diferencias se observan en los primeros tres niveles básicos y los tres superiores.

Asimismo, la autora observa la ausencia casi total del tratamiento de la fraseología de los índices de los manuales, cosa preocupante porque sintomática de la escasa importancia que se le atribuye a la fraseología en los manuales. Por otra parte, Stanič denuncia una incoherencia terminológica en la que se emplean diferentes denominaciones para designar el mismo objeto de estudio o, aún peor, se mezclan teorías y clasificaciones diferentes.

La autora observa, además, que parece haber una voluntad no declarada de los autores de los manuales hacia la simplificación, «depurando» los textos de las UF y, en particular, de aquellas cuyo significado puede resultar más opaco para un hablante no nativo. Por otra parte, no hay coherencia tampoco en la elección de las UF, siendo

estas diferentes en todos los manuales lo que comporta que los estudiantes aprenderán ciertas unidades fraseológicas y no otras en función del manual empleado en el aula.

La académica observa que el papel de la fraseología asume más relevancia conforme sube el nivel de conocimiento. Las colocaciones y las locuciones adquieren más importancia cuantitativa en función del nivel del MCER aunque la calidad de los ejercicios padece los mismos déficits que los de los niveles inferiores: no se trata ni la forma ni el uso de las UF, y se elaboran principalmente ejercicios de enlace en perjuicio de los que permiten practicar el uso en un contexto comunicativo. En este sentido, faltan indicaciones sobre cómo practicar las UF desde un prisma comunicativo, dejando en mano del docente una responsabilidad que, como se ha visto, no es capaz de solventar al no haber recibido formación previa en fraseología.

Tanto Quiroga como Stanič llegan a la misma conclusión. En primer lugar, los manuales tratan una cantidad de UF más elevada en los niveles superiores del MCER que en los inferiores. Desde un punto de vista práctico, se puede afirmar que los autores de manuales de ILE respaldan la corriente fraseodidáctica según la que los niveles superiores son los más adecuados para proporcionar UF.

En este sentido, Lucía López Vázquez (2010) examina exclusivamente cuatro manuales dirigidos a cubrir el nivel C1 de español para extranjeros (ELE) y las conclusiones a las que llega son similares a las que se han destacado previamente para el italiano. La autora esgrime dos características comunes a los manuales de ambos idiomas:

> «primero, el papel secundario que ocupan los ejercicios encaminados a ejercitar la competencia fraseológica productiva del aprendiente, es decir, aquella que le permitiría producir enunciados fraseológicamente adecuados. Y segundo, la ausencia de marcas referidas a las restricciones de uso presentes

> en casi todas las unidades fraseológicas y que resultan fundamentales a la hora de su adquisición». (López Vázquez, 2010: 541)

De la misma opinión es Saracho Arnáiz, quien analiza 16 manuales de ELE en su investigación doctoral y concluye que «los manuales que presentan contenidos fraseológicos más completos y desarrollan actividades fraseodidácticas más completas son los del nivel C», añadiendo que «el profesor de niveles más bajos de ELE no va a encontrar en los manuales de ELE una verdadera ayuda para aplicar al aula de ELE contenidos fraseológicos» (Sarancho Arnáiz, 2015: 926).

Cabe también destacar el análisis Lisciandro elabora para su tesis doctoral (2018) *Las unidades fraseológicas en la enseñanza de una lengua extranjera (italiano - español)*. En su investigación, el autor estudia 25 manuales de italiano como lengua extranjera, pertenecientes a casi todos los niveles que se establecen en el MCER (del A1 al C1), con el objetivo de comprender «qué UF se incluyen en las distintas unidades didácticas, cómo se presentan estas combinaciones léxicas en ejercicios y textos, y con qué finalidad didáctica» (Lisciandro, 2018: 16).

El análisis de los manuales de ILE lleva el investigador a concluir que «el balance de la investigación fraseológica es más que satisfactorio en lo que concierne a la presencia de UFs en los manuales» (Lisciandro, 2018: 218). Con todo, la gráfica propuesta por el autor demuestra como todos los manuales examinados presentan colocaciones y locuciones, y que estas dos esferas de unidades se suelen introducir ya a partir de los niveles elementales e intermedios. Lo mismo ocurre con la que Lisciandro define como *combinazioni semilibere*, es decir, aquellas unidades cuya estructura compuesta por dos sintagmas nominales separados por una preposición en la que, aunque la presencia de un elemento no implique la aparición del otro, la combinación, se considera semilibre por el

empleo obligatorio de la preposición, que varía según el género del segundo sintagma nominal (Lisciandro, 2018: 92).

En conclusión, se podría afirmar que las limitaciones fraseodidácticas descritas para los manuales de italiano no son exclusivas del ILE, sino que se presentan de manera especular también en ELE. Por un lado, los ejercicios no dan suficiente espacio a la producción de dichas UF, limitándose a la primera fase de identificación de los fraseologismos[106] (Kühn, 1992; Ettinger, 2008). En este sentido, desatienden la habilidad activa de producción que el alumnado necesita adquirir para producir UF de manera autónoma puesto que se centran casi exclusivamente en la competencia fraseológica receptiva, es decir, en la habilidad de los aprendientes para descifrar y reconocer estas unidades.

A pesar de que los manuales reserven un papel relevante al universo fraseológico, en ningún momento se hace referencia a posibles contextos de uso, suprimiendo uno de los aspectos principales para la adquisición de las UF. Finalmente, se observa que los manuales prefieren proporcionar material fraseológico en los niveles más altos del MCER, frente a una introducción paulatina ya desde los niveles más elementales.

3.7. Análisis de dos métodos didácticos para impartir fraseología en las clases de ILE

En los últimos años, el estudio de la didáctica de la fraseología italiana ha pasado de ser teórica a práctica y las editoriales han elaborado los primeros manuales exclusivamente dirigidos a impartir fraseología en las clases de italiano como lengua extranjera. Sin embargo, la mayor parte

[106] En una primera fase se activan todos los procesos de descodificación y el alumnado tiene que identificar un fraseologismo dado en un texto auténtico oral o escrito. En la segunda fase, se presentan al discente actividades dirigidas a la comprensión y descifración del significado fraseológico que aparece en una situación comunicativa dada (la fase de reproducción); mientras que en el paso tercero el alumnado debe usar la UF que acaba de aprender en un contexto parecido (la fase de producción) (Kühn, 1992; Ettinger, 2008).

de estas monografías sobre fraseología con intención didáctica[107] se reducen a una mera enumeración de unidades fraseológicas en la que aparecen las UF, generalmente presentadas en orden alfabético y acompañada por su definición. Además, en muchos casos mezclan las tres esferas del universo fraseológico en una amalgama de unidades creando ulterior confusión en los aprendientes y dificultando la labor del profesorado.

Sin intención de elaborar un análisis exhaustivo de todos los manuales específicos de fraseología de ILE, a continuación, se analizan dos de ellos con el objetivo de averiguar su estructura. Los dos libros que se analizan son:

1. Zamora, Pablo; Arianna, Alessandro; Eleonora, Ioppoli; Federica, Simone (2006): *Hai voluto la bicicletta…*, Perugia, Guerra Edizioni.

2. Aprile, Gianluca (2008): *Italiano per modo di dire*, Firenze, Alma Edizioni.

Dichos manuales de fraseología han sido elegidos en función de la calidad de los mismos. En este sentido, se ha querido comparar un manual cuya estructura y contenido se considera de los mejores que se pueden encontrar en el mercado, con uno deficitario y de escaso rigor pragmático.

3.7.1. Zamora, Pablo; Arianna, Alessandro; Eleonora, Ioppoli; Federica, Simone (2006): Hai voluto la bicicletta…, Perugia, Guerra Edizioni

El primer manual que se estudia en profundidad es *Hai voluto la bicicletta…,* cuya autoría corresponde a miembros del ya citado grupo de investigación FRASEMIA de la Universidad de Murcia[108]. Lo primero que salta a la vista de la publicación es su coautoría italo-española, una pista

[107] Véase como ejemplo: Di Natale, Francesco; Nadia, Zacchei (1996): *In bocca al lupo! Espressioni idiomatiche e modi di dire tipici della lingua italiana*, Guerra Edizioni; Andrea Conti (2015), *I 200 Proverbi e Modi di dire italiani più usati: L'Italiano in tasca*, Andrea Conti.

[108] Ver Capítulo II apartado 2.3.1.5. Otras propuestas.

para demostrar que la fraseología italiana se nutre de la española y se impregna de ella. Dicha simbiosis italo-española se observa asimismo en la amplia introducción del manual, en la que los autores proporcionan un detallado marco teórico justificado con clasificaciones y aportaciones de múltiples autores, descritas en el capítulo precedente: Corpas Pastor, Federica Casadei, Alberto Zuluaga, Miriam Voghera, Leonor Ruiz Gurillo, Simonetta Vietri, etc.

De hecho, justifican las razones por las que aparecen tantos hispanistas en un manual de italiano LE afirmando que: «in Italia, al meno fino ad oggi, si è dedicata poca attenzione allo studio della fraseologia e, in particolare, della sua didattica»[109] (Zamora et al., 2006: 8) mientras que en España se ha asistido a un renovado interés por el tema y ha sido tratado en numerosos ensayos, incluso en clave contrastiva entre la lengua italiana y la española.

Aunque se ofrezca una vasta introducción, en la que aparecen aspectos teóricos de la fraseología, el libro se presenta con un declarado fin didáctico, «una proposta di approccio didattico ad un numero limitato di strutture»[110] (Zamora et al., 2006: 7) para que se puedan introducir en el aula ya desde el nivel más elemental. En este sentido, el libro nació de una necesidad práctica encontradas en las aulas, es decir, solventar la percepción de desorientación experimentada por el profesorado a la hora de impartir unidades fraseológicas concretas. El objetivo principal que el texto se propone es permitir al estudiante no solo la descodificación de las UF, sino fomentar la capacidad de reproducirlas autonomentente respetando los valores discursivos y pragmáticos.

La elección de las UF que aparecen en la monografía se rige por criterios discrecionales. Según afirman los autores, al no disponer de estudios

[109] «En Italia, por lo menos hasta el día de hoy, se ha prestado poca atención al estudio de la fraseología y, en particular, de su didáctica». Traducción propia.

[110] «Una propuesta de enfoque didáctico para un número limitado de estructuras». Traducción propia.

científicos sobre la frecuencia de uso, la selección se ha realizado en parte basándose en la bibliografía analizada y en parte en su experiencia personal. Una vez más, frente a una ausencia de literatura académica, se confirma la dificultad a la hora de elegir qué caudal fraseológico impartir en las clases de ILE.

Según la publicación, presentar las UF en un contexto específico o a través de un diálogo permite al estudiante reconstruir el significado y la función pragmática, así como el registro lingüístico en el que se emplea y las características sociolingüísticas de los hablantes que las emplean. Por ello, presentan textos discursivos interactivos sin limitarse a proporcionar enunciados aislados. En ocasiones, dichos textos han sido elaborado *ad hoc* por los autores y, en otras, se han encontrado en Google. De hecho, algunas veces se tiene la percepción de enfrentarse a textos poco reales y que las UF son introducidas de manera poco natural.

Con respecto a la estructura formal, el libro se divide en seis secciones y cada una de ellas a su vez en unidades. Las seis secciones son:

1. Locuzioni «centrali».
2. Locuzioni idiomatiche pragmatiche.
3. Enunciati idiomatici pragmatici ed enunciati pragmatici.
4. Frasi implico-situazionali.
5. Schemi sintattici fraseologici.
6. Segnali discorsivi.

Por lo general, las unidades se rigen mezclando campos conceptuales (ej. los colores, la familia, la cocina, etc.), con funciones comunicativas (ej. para enfatizar, mostrar indiferencia, mostrar incredulidad, mostrar convicción en una afirmación, etc.). Algunas unidades incluyen binomios, metáforas, anáforas, etc. Utilizan, cuando es posible y sobre todo con

respecto a las locuciones, un lexema común[111] puesto que «può aiutare a coglierne il significato e, in alcuni casi, lo sfruttamento in aula di quelle che potremmo definire "metafore universali"»[112] (Zamora et al., 2006: 8)

Finalmente, cada unidad se divide a su vez en tres apartados siguiendo las pautas propuestas por Kühn (1992) y perfeccionadas por Ettinger (2008): reconocer y comprender en una primera fase, memorizar y descifrar en una segunda fase, y, finalmente, producir de manera autónoma las UF en la tercera y última fase. Cabe destacar que no se presenta ningún tipo de marca de nivel que ayude al profesorado a la hora de escoger el ejercicio adecuado. En este sentido, se observa una limitación significativa de tipo práctico puesto que el profesorado no sabe qué ejercicio proponer al alumnado en función del nivel que está cursando.

En conclusión, se puede afirmar que, a pesar de las limitaciones descritas, se trata de un manual bastante completo para trabajar la fraseodidáctica del italiano en el aula por dos razones principales. En primer lugar, por el tipo de metodología que aplica, que considera los aspectos pragmáticos y semánticos que permiten al alumnado utilizar de manera independiente las UF. En segundo lugar, puesto que sus autores son fraseólogos experimentados, se rige en una profunda revisión bibliográfica, destacando su profundo conocimiento del estado del arte en esta materia. En este sentido, el amplio marco teórico ofrecido en la

[111] Con respecto al uso de un lexema común para proporcionar UF en el aula cabe destacar que la opinión de los expertos se divide. Si por un lado Quiroga es favorable, al considerar que «la presencia de un elemento léxico común en las UF que se ha de enseñar es una ayuda para el aprendizaje por parte del estudiante» (2006: 194), por otro lado y como se verá en el siguiente capítulo, hay otros autores como Núñez-Román (2011), Agnieszka Szyndler (2015) o Ruiz Gurillo (2000) que no comparten dicha estrategia. De acuerdo con Agnieszka Szyndler, dicha elección léxica «no nos parece muy eficaz y rentable, dado que dan prioridad al significado compositivo y descartan por completo el significado unitario, esto es, global y figurativo de una locución determinada» (Szyndler, 2015: 211). Asimismo, Ruiz Gurillo también critica este criterio porque no ofrece una visión adecuada de las UF sino que se introducen como si fueran unidades segmentadas en lugar de dejar claro que «siempre se reproducen de manera unitaria y son indisolubles» (Gurillo, 2000: 261).

[112] «Puede ayudar a comprender el significado y, en algunos casos, la explotación en el aula de lo que podríamos llamar "metáforas universales"». Traducción propia.

introducción es de gran utilidad para suplir el déficit formativo del profesorado en cuanto a fraseología.

3.7.2. Aprile, Gianluca (2008): Italiano per modo di dire, Firenze, Alma Edizioni

En los 14 capítulos de *Italiano per modo di dire*, Gianluca Aprile (2008) sigue el criterio de los campos semánticos para organizar su propuesta fraseodidáctica: cuerpo humano, vestimenta, animales, comida y bebidas, etc. En este sentido, la primera limitación que se observa en el manual es que este no explica el criterio escogido para la selección de dichos campos semánticos, y es más, parece ser baladí y no fundada en ningún criterio empírico. Por ejemplo, no se comprende porqué se omiten algunos campos semántico tan productivos en cuanto a fraseología, como el de los colores o el de los sentimientos.

Aprile define en la introducción del manual que su objeto de estudio son los *modi di dire*, denominación excesivamente general dado que puede abarcar varias esferas y que no define realmente lo que se pretende estudiar[113]. En este sentido, ofrece una definición para aclarar qué esfera de la fraseología se pretende examinar, definiendo *modo di dire* como «una frase composta da parole il cui senso non corrisponde alla somma del significato solito dei suoi componenti»[114] (2008: 4). Como se ha visto en el capítulo II, dicha definición es similar a la que Federica Casadei (1995) proporciona para describir las locuciones, es decir, unidades «il cui significato non è composizionale, cioè non è funzione dei significati dei componenti»[115] (1995: 28). Aprile, sin embargo, no emplea ni el término *fraseologia* ni *unità fraseologiche*.

[113] Cabe recordar que, como se ha visto en el capítulo 2.3. La fraseología en la lingüística italiana, el término «modi di dire» es demasiado general y Faloppa (2011) en el *Dizionario Treccan*i propone denominar estos compuestos como «espressioni idiomatiche».

[114] «Una frase compuesta por palabras cuyo significado no corresponde a la suma del significado habitual de sus componentes». Traducción propia.

[115] «Cuyo significado no es compositivo, es decir, no es relacionado con los significados de los componentes». Traducción propia.

Por otra parte, el manual de Aprile proporciona una definición con el objetivo de distinguir un *modo di dire* de *proverbios*. En este sentido, afirma que:

> «nel proverbio è sempre sottinteso un principio didattico e morale, un avvertimento, un consiglio o una massima che hanno origine dall'esperienza. Per questo il proverbio è chiamato anche «sapienza dei popoli». A differenza del proverbio, il modo di dire è nella maggior parte delle volte un «paragone accorciato», una metafora semplificata, per esempio: essere lento come una lumaca diventa essere una lumaca»[116]. (Aprile, 2008: 4)

Dicha distinción no tiene ningún precedente entre los estudios de fraseología italiana o española. Además, y como apunta Saracho Arnáiz, «el hecho de que aún hoy siguen vigentes nomenclaturas variadas como «frase hecha», «dichos», «idiomatismos», «giros», etc. en manuales u otro tipo de obras nos da una idea de la falta de claridad y de fundamento fraseológicos que muchos manuales transmiten» (2015: 922).

El enfoque didáctico que el manual pretende proporcionar tampoco se clarifica en la introducción. En cambio, es fácil comprender que se trata principalmente de ejercicios de tipo *cloze*, tipología que deja al alumnado la posibilidad de elegir entre varias propuestas para rellenar un hueco, y de ejercicios en los que el estudiante tiene que unir una UF a su paráfrasis o definición. No obstante, dichos ejercicios no son dirigidos hacia la producción autónoma de las unidades por parte del alumnado, sino que se centran sobre todo en su descifración y memorización.

Por otra parte, los ejercicios son escasamente comunicativos y no fortalecen ni la comprensión oral ni escrita de dichas unidades, puesto

[116] «El proverbio siempre implica un principio didáctico y moral, una advertencia, un consejo o una máxima originada en la experiencia. Por ello, el proverbio también se llama «sabiduría de los pueblos». A diferencia del proverbio, el dicho es en la mayoría de los casos una «comparación abreviada», una metáfora simplificada, por ejemplo: ser tan lento como un caracol se convierte en ser un caracol.». Traducción propia.

que parecen no proporcionar material original, sino breves frases elaboradas *ad hoc* con las que trabajar. En este sentido, el estudiantado no puede superar las principales problemáticas que presentan las UF, es decir, reconocer su nivel pragmático y semántico, así como el registro lingüístico en el que los hablantes la insertan y utilizan.

Analizando los niveles del MCER que el manual pretende abordar, lo primero que deja patente ya en su portada es a quién se dirige dicha obra. En este sentido, abarca un espectro muy amplio, del nivel A2 hasta el C1 de ILE. Las razones por las que se ha dejado de lado tanto el nivel A1 y el C2 no son razonadas en la introducción. Sin embargo, un rápido estudio de los ejercicios podrían verificar una posible hipótesis.

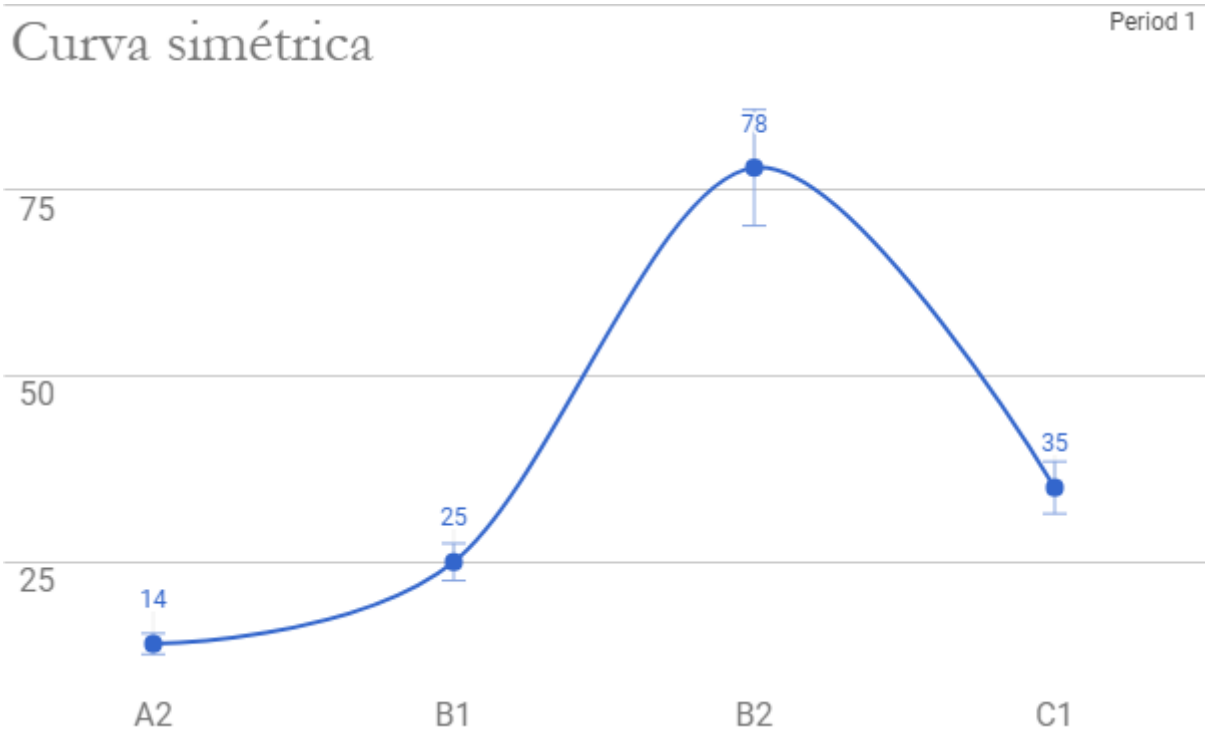
Cada ejercicio lleva en su margen superior derecho una marca para guiar el profesorado: los ejercicios que presentan una bola están pensados para el nivel A2, dos bolas para el B1, tres bolas para el B2 y cuatro bolas para estudiantes del nivel C1. Cabe destacar que Aprile no explica, ni se consiguen intuir de manera autónoma, los criterios según los cuales se ha decidido qué nivel atribuir a cada ejercicio.

A través de un rápido cálculo se observa que, de un total de 152 ejercicios, sin considerar los crucigramas con los que termina cada unidad ya que estos no poseen ninguna marca de nivel, para el nivel A2 el autor ha elaborado tan solo 14 ejercicios. Sin embargo, si se examina el nivel B1 se observa que el número de ejercicios se eleva a 25 y en B2 a 78 ejercicios. Finalmente, el número de ejercicios presentados el nivel C1 desciende nuevamente a 35.

Por lo tanto, se puede observar que más del 50% de los ejercicios del libro han sido elaborados pensando a un nivel B2, mientras que, en los demás niveles, sobre todo los más elementales, se proporciona un número muy escaso. La cantidad, por ende, de ejercicios proporcionados forma una curva simétrica (véase Figura 5) que tiene su pico en el nivel B2 y

desciende drásticamente hacia los otros extremos, el nivel A2 por un lado y el C1 por el otro.

Figura 5. Curva simétrica de ejercicios por niveles del MCER en el manual de Aprile (2008).



Fuente: elaboración propria.

Tras este análisis se podría deducir la razón por la que el autor excluye de su manual el nivel A1 y el C2. En este sentido, parece que Aprile considere que el universo fraseológico debería de ser presentado en aula solo a partir del nivel A2 y sin proporcionar demasiadas estructuras; progresivamente, habría que incrementar esta cantidad hasta el B2, nivel en el que se impartirá la mayor parte de las UF y, desde este, reducir significativamente el número hasta el C1. Finalmente, cabe resaltar que el autor considera que el nivel A1 y el C2 son inadecuados para aprender fraseología, a la vista de los resultados de la investigación.

En conclusión, el manual analizado no parece ser lo suficientemente estructurado como para poder considerarlo fiable a la hora de utilizarlo en aula. Por un lado, carece significativamente de un marco teórico en

el que basarse, incluso destaca la falta de referencias del autor a la revisión bibliográfica previa en fraseología necesaria para la creación de este tipo de materiales didácticos. Por otra parte, de acuerdo con Serradilla Castaño, «no creemos que dar listas de expresiones y hacer únicamente ejercicios de llenar huecos sea la mejor solución» (2000: 658). Así, se ha observado que los ejercicios no son comunicativos, no incluyen material real y se limitan a la descodificación y memorización de las UF, olvidando que el logro final de la fraseodidáctica es la producción autónoma de la UF por parte del alumnado en un registro y contexto determinado.

En este sentido, las unidades fraseológicas de forma aislada representan una manera improductiva de enseñanza-aprendizaje. A pesar de que se atribuya un nivel de referencia a cada ejercicio, no solo no se explican los criterios elegidos para dicha atribución, sino que su distribución no es homogénea, ya que se concentra en gran mayoría en el nivel B2 dejando completamente de lado el A1 y el C2, sin proporcionar alguna explicación respecto a una elección tan significativa.

3.8. La fraseodidáctica italiana y española a través de los diccionarios

«Prácticamente, desde los inicios de la lengua encontramos recopilaciones que han tratado de atrapar la creatividad y el ingenio de los hablantes» (Castillo Carballo, 2017: 85). Por ello, para hablar de fraseología es necesario mencionar la fraseografía. Término ya empleado por Tristá Pérez (1998), la fraseografía es una una rama de la fraseología cuyo objetivo es determinar los problemas teóricos y prácticos que plantea la confección de diccionarios fraseológicos.

Corpas Pastor y Morvay, al observar que los principales diccionarios del español «ignoraban completamente la palabra fraseografía» (Corpas Pastor y Morvay, 2002: 166) proponen una definición propia del término definiéndolo como «disciplina que se dedica a elaborar y poner en

práctica los principios de tratamiento de las unidades fraseológicas en los diccionarios generales y fraseológicos» (2002: 166).

No es intención de este trabajo doctoral estudiar pormenorizadamente el tratamiento que los diccionarios proporcionan a las unidades fraseológicas. A pesar de ello, las obras lexicográficas desempeñan un importante valor didáctico añadido (Marello, 1993) y representan un valioso instrumento para enseñar/aprender una lengua extranjera (Dubois, 1970; Santamaría Pérez, 2000), por lo que se considera necesario ofrecer unas pinceladas de las principales aportaciones fraseográficas en los diccionarios monolingües italianos y españoles, así como bilingües italiano-español.

3.8.1. Diccionarios monolingües españoles

La fraseología en los diccionarios monolingüe españoles ha sido, por lo general, ampliamente abordada y se dispone de materiales bien estructurados que los que se basan los avances teóricos planteados por la academia. Uno de los primeros diccionarios fraseológicos realizados en español es el de Manuel Seco (2004). Su *Diccionario fraseológico documentado del español actual* es «el más sistemático y completo, tanto a nivel de cantidad de secuencias como de exhaustividad en su presentación y definición (Alessandro, 2011: 251) y el periódico El País lo definió como «el más ambicioso en su campo» (El País, 12.10.2004).

La publicación contiene más de 16 mil unidades fraseológicas extraídas de la prensa, ensayos o novelas y, por varias entradas, se proporciona también la motivación de la expresión. Según Seco (2004), el diccionario es destinado a todo el que sepa leer español, pero es su propio autor quien le atribuye en primer lugar una finalidad didáctica afirmando que «es muy útil para los profesores de español y los alumnos avanzados que lo aprenden»[117]. El diccionario recopila en su gran mayoría locuciones,

[117] El País, 12.10.2004. Véase <https://elpais.com/diario/2004/11/12/cultura/1100214002_850215.html>.

colocaciones y paremias, y Simone (2011) destaca que se trata de una de las pocas herramientas fraseológicas que proporciona cierta atención a las unidades fraseológicas pragmáticas (UFP).

Otros valiosos instrumentos fraseográficos en lengua española son los tres diccionarios fraseológicos de Isabel Penadés Martínez. En particular, el primer volumen atañe a las locuciones verbales (2002), el segundo sobre las locuciones adverbiales (2005) y, el último aborda las locuciones nominales, adjetivas y pronominales (2008). Se trata de trabajos detallistas y bien acabados en los que se percibe la intención didáctica de la autora.

De hecho, cada locución tiene asignado un nivel de referencia del MCER aunque, no aparecen los niveles A1, A2 y el C2, en línea con la corriente que considera la fraseología como un tema propio de los niveles intermedios (Ruiz Gurillo, 2001; Koike (2001); López-Jiménez (2013: 338); etc.). Por otra parte, las entradas identifican el registro de la locución y ofrecen un amplio espectro de uso recopilados del Corpus de referencia del Español Actual (CREA), tanto de la lengua oral como de la escrita. Por último, se ofrecen también antónimos y sinónimos de la locución e información gramatical y pragmática para facilitar la comprensión de un estudiante de ELE.

3.8.2. Diccionarios monolingües italianos

Paula Quiroga (2006) realiza un profundo y exhaustivo análisis del tratamiento de la fraseología en los principales diccionarios monolingües italianos[118] reconociendo que, a pesar de que los lexicógrafos de este idioma «prestan bastante atención» a los fenómenos fraseológicos, «no se ofrece un tratamiento sistemático de estas unidades» (2006: 135). De las cuatro obras analizadas por la autora[119], la única que se ocupa de

[118] Duro (1986-1994), De Mauro (1999; 2000), Zingarelli (2003), Sabatini y Coletti (2003).

[119] Consultable gratuitamente en la versión en línea a la dirección del *Corriere della Sera*: <http://dizionari.corriere.it/cgi-bin/sabcol/trova>

todas las UF en general es el Sabatini y Coletti (2003) que resulta ser la publicación más completa en cuanto al tratamiento del universo fraseológico. El trabajo de De Mauro (1999, 2000) viene, según la autora, en segundo lugar en cuanto al tratamiento fraseográfico.

En el capítulo II de la presente tesis doctoral se ha descrito pormenorizadamente el tratamiento que la fraseología recibe en el *Grande dizionario italiano dell'uso* (GRADIT) de De Mauro (1999), así como en *Il dizionario della lingua italiana per il terzo millennio*, que el mismo autor elabora en el 2000 a partir del GRADIT, incluyendo más de 30.000 *polirematiche* como entradas independientes.

Estudiando la producción de diccionarios monolingües italianos específicos sobre fraseología, es decir, diccionarios fraseológicos, se puede comprobar que, a diferencia del español, la producción es más bien escasa, por no decir ínfima. Una de las primeras monografías creadas con declarados fines didácticos es la de Giuseppe Pittáno, titulada *Frase fatta capo ha. Dizionario dei modi di dire, proverbi e locuzioni di italiano* (2009). Casi 1400 unidades representan el corpus de esta obra en la que el autor organiza, en orden alfabético, los que llama *modi di dire*, los proverbios y las locuciones[120].

Según la clasificación que se emplea en la presente tesis doctoral, el concepto del *modo di dire* propuesto por el autor, y recuperado de Aprile (2008: 4) es asimilable a la clase de las locuciones adjetivas sobre la base de comparaciones estereotipadas. Ya en el título del diccionario, el autor declara querer incluir también las locuciones aunque, desafortunadamente, no ofrece una definición concreta para dicha clase de unidades. En este sentido, Pittáno se limita a utilizar el término como análogo de modi di dire «la combinazione di più parole il cui senso

[120] La definición de Pittano (2009: 3) para *modi di dire* y *proverbi* es homóloga a la que proporciona Aprile (2008: 4).

non corrisponde alla forma del significato normale dei loro componenti»[121] (2009: 4).

Las entradas de las unidades se presentan en orden alfabético en función de la primera palabra que la componen y se menciona por cada una el significado principal y el origen etimológico, es decir, la motivación de la UF. En numerosas ocasiones se presentan también sinónimos o variantes estructurales[122] tratadas en la misma entrada por practicidad, según afirma el autor en su introducción.

Otra interesante monografía que cabe destacar en este apartado es la de Monica Quartu y Elena Rossi: *Dizionario dei modi di dire* (2012). Se trata de un diccionario monolingüe que recopila buena parte de la fraseología italiana en un único diccionario consultable gratuitamente y en línea en la página del periódico *Il Corriere della Sera*[123]. Con sus 528 páginas y sus más de ocho mil entradas, se trata de la primera obra fraseográfica de gran interés didáctico, dado que recopila distintas unidades fraseológicas de diversas áreas del conocimiento[124].

El diccionario se presenta bien estructurado, de manera que el alumnado puede utilizarlo por su cuenta, sin el acompañamiento del profesorado, introduciendo en el buscador en línea un lema y jugando a analizar los resultados. En cambio, desde un punto de vista taxonómico, y como el mismo nombre de la obra deja intuir, se trata de un verdadero *popurrí* fraseológico en el que se entremezclan frases proverbiales, citas, locuciones de todo tipo y fórmulas rutinarias, sin que se ofrezca ninguna pista sobre su clasificación. Por todo ello, se puede afirmar que, a pesar de ser un instrumento muy valioso desde un punto de vista didáctico, no consigue solventar los problemas terminológicos y taxonómicos que se

[121] «La combinación de varias palabras cuyo significado no corresponde a la forma del significado habitual de sus componentes». Traducción propia.

[122] Para más información véase: Corpas Pastor (1996: 28-30).

[123]Consultable gratuitamente en su versión en línea a la dirección del Corriere della Sera <http://dizionari.corriere.it/dizionario-modi-di-dire/>.

[124] Política, periodismo, cine y televisión, matemáticas, etc.

han evidenciado en los apartados anteriores con respecto a la lengua italiana.

Solo quedan excluidas de este acervo fraseológico las colocaciones, esfera a la que, en cambio, Paola Tiberi (2012) dedica un diccionario específico: el *Dizionario delle collocazioni. Le combinazioni delle parole in italiano*. La autora define las *collocazioni* como «espressioni formate da due o più parole che per uso e consuetudine lessicale formano una unità fraseologica non fissa ma riconoscibile»[125] (Tiberi, 2012: 3) y no es una casualidad que, en la portada del mismo diccionario, aparezcan unas piezas Lego® encajadas, puesto que, como se ha descrito, los componentes de dichas UF poseen una fuerte atracción sintáctica.

El *Dizionario delle collocazioni* estudia más de 200 mil colocaciones en más de seis mil entradas entre adjetivos, sustantivos y verbos. Cada lema proporciona al lector una amplia elección de las colocaciones más significativas en función del uso y de la categoría gramatical que desempeñan. Cabe destacar que Tiberi ha eliminado de su corpus las colocaciones raras y específicas y las que son demasiado generales[126]. De particular interés desde un punto de vista didáctico es la puntualización que la autora hace con respecto a la posición entre colocativo y colocado. En este sentido, para estudiante de ILE puede ser útil saber si la posición de los componentes de una colocación es fija o si se puede alterar el orden, posponiéndolos o anteponiéndolos, sin que el valor pragmático de la colocación sufra variaciones significativas.

Con todo, el diccionario resulta bastante completo y valioso, tanto para el profesorado/alumnado como para quien se dedique a la investigación. Sin embargo, cabe destacar que la autora decidió tomar

[125] «Expresiones compuestas por dos o más palabras que por su uso y costumbre léxica forman una unidad fraseológica que no es fija sino reconocible» Traducción propia.

[126] En este sentido, la autora ha eliminado los verbos avere y essere, así como todos los adjetivos excesivamente banales como *buono, cattivo, bello, brutto*, etc., siempre y cuando dichos adjetivos no formen un colocativo fuerte o reconocibles, como *essere in cammino, buon appetito*, etc.

en consideración solo las colocaciones con sustantivos en número singular, eliminando *a priori* todos los declinados en plural, lo que supone una lacuna lexicográfica.

3.8.3. Diccionarios bilingües italiano-español

Distintos lexicógrafos y fraseólogos subrayan las limitaciones que presenta la lexicografía bilingüe italiana-español. De acuerdo con Marello (1989: 194), esta lexicografía bilingüe ha sido, y en cierto modo sigue siendo hoy en día, una de las más pobres y menos desarrolladas. En su análisis sobre el tratamiento de la fraseología en los diccionarios bilingües italiano-español, Paula Quiroga (1996) pone de relieve una situación caótica con respecto al tratamiento de la fraseología. Alessandro llega incluso a afirmar que dichos estudios sufren «ciertas carencias de rigor en la adopción de criterios lexicográficos» (Alessandro, 2011: 250). Y, finalmente, Carmen Navarro (2005) opina que la lexicografía bilingüe italiana-española presenta limitaciones profundas en varios aspectos. En un estudio en el que examina dos diccionarios bilingües español-italiano,[127] concluye que el tratamiento de la fraseología en ambos diccionarios se aborda de manera confusa y escasa.

A pesar de estas limitaciones descritas por las autoras citadas, en este apartado se pretende destacar como ejemplo de buenas prácticas el *Diccionario Hoepli italiano-español y español-italiano*, obra lexicográfica elaborada por Laura Tam (1997)[128], profesora en la Università Bocconi, y publicado por Hoepli. Con más de 450 mil entradas, 130 mil lemas, 70 mil términos técnicos y 2.500 americanismos se convierte, con toda probabilidad, en el diccionario de referencia para dicho par de lenguas. Asimismo, la publicación proporciona más de 70 mil locuciones y Quiroga considera que esta obra representa:

[127] Tam (1997; 2004 2.ª ed.) y Sañé-Schepisi (2005).

[128] Consultable gratuitamente en la versión en línea: <http://www.grandidizionari.it/Dizionario_Italiano-Spagnolo.aspx?idD=4>

> El único diccionario[129] que reserva un espacio específico para la fraseología. Al final del artículo, tras un punto y aparte, introduce todas las UF [Unidades Fraseológicas] mediante un riguroso orden alfabético. Este párrafo está precedido por la abreviatura FRAS. Que, como señala la autora en el apartado de signos gráficos de las hojas preliminares, indica que a partir de ahí se introduce la fraseología italiana. (Quiroga, 2006: 85)

Sebastián Carbonell es autor del *Dizionario fraseológico completo italiano-spagnolo e spagnolo-italiano* (1950-1953), editado también por Hoepli. El diccionario, que ofrece más de 80 mil entre vocablos y modi di dire, según la nomenclatura empleada por el autor, representa un trabajo pionero en el ámbito de la lexicografía contrastiva entre este par de lenguas ya que, por primera vez, se presenta un corpus extenso de unidades fraseológicas italianas y españolas de manera contrastiva. La obra refleja un gran número de locuciones y frases proverbiales teniendo a menudo en cuenta las variantes latinoamericanas.

Entre el final de los ochenta y el principio de los noventa del siglo pasado, ve la luz una versión en dos tomos actualizada de dicho diccionario. En primer lugar, se edita el *Dizionario fraseologico completo italiano-spagnolo e spagnolo-italiano. Parte spagnola-italiana* (1989) y, el año sucesivo, se presenta la segunda parte, el *Dizionario fraseologico completo italiano-spagnolo e spagnolo-italiano. Parte italiana-spagnola* (1990). En más de 2.400 páginas, la publicación registra 165 mil entradas y cerca de 80 mil unidades fraseológicas en la parte italiana-española y 85 mil en la parte española-italiana. En efecto, ofrece una gran cantidad de locuciones y refranes de los dos idiomas objeto de estudio, incluyendo

[129]Los diccionarios revisados por Quiroga (2006) es su estudio son: 1) Ambruzzi L., *Nuevo dizionario spagnolo-italiano, italiano-spagnolo*. Torino, Paravia, 7.ª ed. 1973 (1.ª ed. 1949); 2. Carbonell S, D*izionario fraseologico italiano-spagnolo, spagnolo-italiano*. Milano, Hoepli, 1997 (1.ª ed. 1950-1953); 3. AA. VV., *Diccionario avanzado italiano Zanichelli. Italiano-spagnolo, español-italiano*. Barcelona, Vox, 1980; 4) Gallina A. M., *Dizionario spagnolo-italiano, italiano-spagnol*o. Milano, Mursia, 7.ª ed.1997 (1.ª ed. 1990); 5) Calvo Rigual C. y Giordano Gramegna A., *Diccionario italiano-español español-italiano.* Barcelona, Herder, 1995; 6) Tam (1997).

terminología específica de distintas áreas del conocimiento: de carácter artístico, geográfico, jurídico, marítimo, literario, etc.

Otra recopilación fraseográfica elaborada desde una perspectiva contrastiva italiana-española es la de Pablo Zamora, italianista de la Universidad de Murcia. El diccionario *Spagnolo-italiano: espressioni idiomatiche e proverbi* (1997), fruto de cinco años de investigación, es de particular interés para la presente tesis doctoral, puesto que ha sido elaborado con una declarada intención didáctica, es decir, para hablantes de italiano como lengua materna que aprenden español como LE. Como afirma el autor, la obra es una «recopilación de expresiones idiomáticas, refranes y frases pragmáticas española e italianas más frecuentes» (Zamora, 1997: 1) y tiene como objetivo:

> [...] cubrir en la medida de lo posible, el vacío teórico que los diccionarios bilingües español-italiano presentan respecto a las equivalencias lingüísticas en ambas lenguas, constituyendo un útil apéndice que facilite el aprendizaje de esa área semántica tan completa pero, a su vez, imprescindible, dado el uso masivo que se hace de ellas en la lengua coloquial. (Zamora, 1997: 1)

Para seleccionar el corpus que constituye la obra, el autor sigue el criterio de la frecuencia de uso, recogiendo las UF propias de caudal de un hablante medio. La estructura formal del diccionario es peculiar con respecto a un diccionario tradicional, puesto que las UF no son clasificadas en función del orden alfabético del primer componente de la unidad, sino se rigen según áreas conceptuales (amenaza, ambición, miedo, verdad, etc.) y dentro de este, por orden alfabético. De modo que, dentro del área conceptual, «abatimiento» aparece en orden las unidades *caérsele el alma a los piés, estar bajo de moral* y *estar bajo mínimos*, etc. Cada entrada es acompañada por su definición en español parafraseada, un ejemplo de uso en contexto y su correspondiente traducción al italiano en cursiva.

En conclusión, se puede afirmar que, a pesar de que los diccionarios monolingües y bilingües italiano-español sean herramientas útiles en la didáctica del léxico, en lo que concierne a la fraseología, queda todavía camino por recorrer. En particular, como demuestran las tres obras fraseográficas de Penadés Martínez (2002, 2005 y 2008), se ha observado que la lengua española dispone de diccionarios fraseológicos monolingües más completos y sistemáticos que la italiana, y son más ricas en cuanto a UF presentadas.

Con respecto a los diccionarios fraseológicos bilingües italiano-español, se demuestra una escasez de materiales específicos para el estudio contrastivo de la fraseología en el par de lenguas objeto de estudio. Dicha laguna de materiales no puede que repercutir negativamente en la labor diaria del profesorado y, en general, en la didáctica de la fraseología en las clases de ILE.

CAPÍTULO IV

INDICADORES FRASEODIDÁCTICOS Y APLICACIONES PRÁCTICAS

4.1. Introducción

> A menudo es difícil encontrar un eje unificador que contenga un cierto número de expresiones con las que los alumnos puedan trabajar, sin que crean que se trata de fenómenos totalmente caóticos de la lengua que están estudiando (Quiroga, 1996: 54)

A la hora de establecer qué unidades fraseológicas enseñar en cada nivel del MCER, se hace necesario analizar los principales criterios unificadores propuestos por la comunidad académica. Cabe destacar que, a pesar de que se haya planteado desde diferentes estudios[130] la necesidad de establecer criterios claros y empíricos para que el profesorado pueda disponer de indicadores fiables con la finalidad de proponer al alumnado de ILE actividades didácticas sobre fraseología, aún no se dispone de los medios necesarios para alcanzar dicha meta.

Como se ha examinado en el capítulo anterior, las aportaciones fraseodidácticas son cada vez más numerosas y, en los últimos años, se está profundizando en dicha rama aplicada de la fraseología. Dicho interés ha llevado a la academia a elaborar diversos instrumentos[131] que

[130] Entre ellos se destaca Alessandro, 2011; 2015; Saracho Arnáiz, 2015, Daša Stanič, 2014; Núñez-Román, 2011; Molina García, 2006; Agathe Rakotojoelimaria, 2004; Penadés Martínez, 2002, 2005, 2008; Marta Higuera, 1997; etc.

[131] Entre dichos instrumentos destaca la elaboración de diccionarios fraseológicos con fines didácticos (véase Agathe Rakotojoelimaria, 2004) y la elaboración de corpus fraseológicos ad hoc que, combinados

permiten al profesorado abordar el universo fraseológico en sus clases, con el fin de colmar el vacío estructural que se ha descrito a lo largo del presente trabajo doctoral. No obstante, la revisión bibliográfica también ha evidenciado que la elección de las UF que componen muchos de los corpus fraseológicos es prevalentemente subjetiva y basada en la propia experiencia subjetiva docente.

Esta situación es debida a dos factores principales. En primer lugar, la escasez de material de calidad que guíe el profesorado en la enseñanza de la fraseología. A pesar de que se dispone de materiales fraseológicos elaborados con fines didácticos, se comparte la opinión de Alessandro (2011: 225) al afirmar que «en el ámbito de la enseñanza de E/LE y, de forma aún más notable, de I/LS, la oferta de este material es bastante reducida». El segundo factor, y tal vez el principal, es la escasez de corpus fraseológicos cónsonos con las últimas investigaciones el universo fraseológico. De acuerdo con Agathe Rakotojoelimaria:

> Aunque existen muchos repertorios monolingües dedicados de manera específica a las unidades fraseológicas españolas, la mayoría de ellos incluyen distintas clases de estas unidades mezcladas, sin preocuparse de las posibles clasificaciones que se les puedan aplicar. (Rakotojoelimaria, 2004: 214-215)

La ausencia de rigor científico a la hora de compilar un corpus fraseológico plantea la necesidad de establecer y delimitar los que en este trabajo se definen como *indicadores fraseodidácticos* (IF), es decir, un conjunto de herramientas diagnósticas, tanto cualitativas como cuantitativas, cuya aplicación permite didactizar la fraseología en función del MCER desde un prisma didáctico. En este sentido, por un lado evalúan y describen y, por otro, ejecutan la didáctica de fraseología. Los IF permiten medir empíricamente el comportamiento de la fraseología de un

con una propuesta educativa, plantean una determinada metodología docente (véase Alessandro, 2011; 2015).

idioma en función de algún(os) criterio(s) lingüísticos y se nutren de las propuestas didácticas y de los estudios que han precedido este trabajo. Por último, se dirigen en manera particular al profesorado de LE para que éste los incorpore a sus procesos de toma de decisiones para la elaboración de materiales fraseodidácticos y la asignación de una UF a un determinado nivel del MCER.

Con el objetivo de definir los principales indicadores fraseodidácticos, se realizará una profunda revisión bibliográfica de los principales criterios propuestos para proporcionar fraseología al alumnado en el presente capítulo. En primer lugar, se elaborará un análisis crítico de los indicadores fraseodidácticos más productivos para organizar el material fraseológicas. Dicho examen dejará en evidencia exclusivamente los criterios empíricos omitiendo los que padecen subjetividad, es decir, los que se basan exclusivamente en la experiencia personal de su autor.

4.2. Agrupaciones habituales de los IF

Impartir fraseología en el aula significa actuar, por lo menos, en tres fases didácticas secuenciales: en primer lugar, es necesario plantearse qué criterios formales seguir para seleccionar, dentro del vasto universo fraseológico, la fraseología a enseñar. En segundo lugar, se debe dirigir la mirada hacia el alumnado y, en particular, a los niveles del MCER más adecuados para impartir una determinada unidad[132]. En la última fase, se confeccionan ejercicios específicos a través de los cuales el alumnado podrá reconocer, memorizar y utilizar la fraseología escogida. Este proceso dinámico genera una labor que necesita una gran inversión de tiempo del que, en la mayoría de los casos, el profesorado no dispone.

[132] Como se ha visto en el capítulo II, el presente trabajo parte de la base de que, para alcanzar una competencia fraseológica activa de la LE, es necesario que el alumnado siga un formación continua y progresiva, que significa una puesta en marcha precoz de un planteamiento didáctico específico para el tratamiento de estas unidades en el aula a partir de los niveles más básicos del MCER.

Por otra parte, se ha demostrado133 que el profesorado de ILE por un lado no posee adecuada formación para llevar a cabo dicha tarea plenamente y, por otra parte, desconoce los pocos materiales fraseodidácticos que abordan la fraseología italiana. En este sentido, el docente juega un papel central y casi exclusivo en la selección de las UF lo que inevitablemente conlleva un tratamiento subjetivo y aleatorio de la fraseología del ILE (Zuccalà, 2018).

El profesorado que percibe la importancia de impartir fraseología en el aula se ve obligado a encargarse de establecer autónomamente cuáles UF impartir en su clase en función del nivel y las exigencias de su alumnado (Alessandro, 2011). Confiando en su dominio de la LO, se deja guiar por las unidades didácticas de los manuales o por la conversación natural con el alumnado, lo que implica una selección desestructurada de la fraseología.

Por todo ello, es necesario disponer de criterios empíricos para identificar los que en este trabajo se han definido como indicadores fraseodidácticos. A continuación, el estudio se centrará en el análisis crítico de los indicadores fraseodidácticos más productivos propuestos por la comunidad científica para agrupar la fraseología en función de los niveles del MCER.

4.2.1. Léxico clave

En el capítulo seis del Marco Común Europeo de Referencia para las lenguas: aprendizaje, enseñanza y evaluación (MCER, 2001) se sugieren unas pautas para la selección del léxico, tarea obligatoria para quienes se aproximan a elaborar materiales de exámenes y manuales para una LE[134]. En particular, el MCER propone elegir palabras y frases claves en función de las «áreas temáticas necesarias para la consecución de tareas comunicativas adecuadas a las necesidades de los alumnos - y

[133] Véase Zuccalà (2018).

[134] A este respecto, véase también Villarini (2017 y 2021).

añade que dichas áreas tienen que encarnar - la diferencia cultural y, en su caso, los valores y creencias importantes compartidos por el grupo o los grupos sociales cuya lengua se está enseñando» (Consejo de Europa, 2001: 6.4.7.3.). Sin embargo, cabe observar como dicha propuesta da prioridad al significante con respecto al significado.

Como se ha hecho referencia previamente, la fraseología y el léxico comparten aspectos lingüísticos y, de hecho, Daniel Molina García (2006) pública una interesante monografía titulada *Fraseología bilingüe. Un enfoque lexicográfico pedagógico* en este sentido. Refiriéndose de manera específica a las locuciones, identifica unos criterios unificadores[135] para elaborar una colección de UF y el primero que propone es la agrupación por palabras claves.

Con ello el autor plantea la compilación de un corpus fraseológico en función de un elemento léxico común. En este sentido, el significante común *«testa»* (cabeza) podría representar el elemento unificador de un posible listado de UF italianas entre las que aparecen unidades como: *Testa o croce*, es decir, sortear algo al azar lanzando en el aire una moneda; *Avere la testa tra le nuvole*, para identificar a una persona distraída y fuera de la realidad cotidiana; y *Essere in testa*, estar delante de todos los demás o ser el primero de una lunga serie.

En dichas UF, el significante común *testa* asume un significado diferente en cada una de las unidades: en *Testa o croce* se refiere a la cara de la moneda; en *Avere la testa tra le nuvole* se refiere metafóricamente a la parte del cuerpo y en *Essere in testa* se refiere a la parte más adelantada de una fila o formación. A pesar de que el elemento unificador tenga distintos significados, esta tipología de agrupación de UF es la más utilizada en los manuales y en los glosarios fraseológicos elaborados con

[135] Molina García (2006: 200-206): 1. Agrupación por palabra clave; 2. Agrupación por temas; 3. Agrupación por presentación alfabética; 4. Agrupación por "Metáfora conceptual"; 5. Agrupación por estructura gramatical; 6. Agrupación por contextos situacionales; 7. Agrupación por función o concepto.

fines didácticos[136] en los que se presentan alfabéticamente agrupadas por palabras claves.

De acuerdo con la agrupación en función de dicho indicador fraseodidáctico, Alarcón Collado y López García (1992) proponen contextualizar las unidades a través de fragmentos de la lengua hablada o escrita, y la puesta en práctica de dichas unidades en distintos niveles de competencia a través de las cuatro destrezas, oral, auditiva, escrita y lectora. Entre los ejercicios, las autoras proponen una serie de actividades basadas en esquemas gráficos en los que las UF son agrupadas alrededor de una palabra clave común. Así, en función de la palabra «cabeza», presentan varias las locuciones como: *estar mal de la cabeza, darse con la cabeza contra la pared, no caber en la cabeza, calentarse uno la cabeza*, etc. Una vez más, al reunir UF bajo un lema particular, se da prioridad al significante que prevalece sobre el significado compositivo de la unidad fraseológica.

Otra autora que propone emplear dicho IF para didactizar el universo fraseológico en las clases de ELE es Forment Fernández (1998). En un artículo sobre la didáctica de la fraseología que ya se ha convertido en un clásico, la autora presenta una propuesta para trabajar en el aula las locuciones somáticas y basa su metodología en la elección de un lexema común que identifique una parte del cuerpo humano interna o externa. En este sentido, establece un subconjunto con las locuciones que incluyen los vocablos «ojo», «cabeza», «pie» y «boca». Como resultado, la investigadora obtiene un listado de locuciones somáticas al que aplica un segundo criterio con el objetivo de explotarlo en el aula de ELE: las metáforas universales (Lakoff & Johnson, 1991), concepto que se analizará en profundidad a lo largo de este trabajo doctoral. Con todo,

[136] Véase como ejemplos: Di Natale, Francesco; Nadia, Zacchei (1996): *In bocca al lupo! Espressioni idiomatiche e modi di dire tipici della lingua italiana*, Guerra Edizioni; Andrea Conti (2015), *I 200 Proverbi e Modi di dire italiani più usati: L'Italiano in tasca*.

en este apartado se focaliza en cómo la autora propone centrar la elección de la fraseología en el IF de las palabras claves.

A pesar de que haya una parte de la comunidad científica que sugiere el empleo de dicho IF a la hora de didactizar la fraseología, cabe destacar que existe una corriente que lo rechaza o, mejor dicho, lo considera insuficiente y en algunos casos limitativo. A pesar de que no se puede cuestionar la eficacia de dicho sistema (Alessandro, 2011), cabe apuntar que, como apunta Saussure (1915) y luego Chomsky (1965, 1980), la lengua no es una mera enumeración de vocablos, sino que posee una estructura propia y, tratándose del universo fraseológico esta estructura no es solo sintáctica, sino pragmática.

Las palabras no se asimilan de manera aislada (Baralo, 2006), sino que para que se inserten en el lexicón de un hablante de LE deben establecerse redes semánticas, y esta premisa es aún más acertada si se habla de unidades fraseológicas. La fraseología no puede ser tratada desde la misma perspectiva que el léxico puesto que, a pesar de las similitudes que presentan, es el significado unitario de los componentes de una UF lo que las convierte en una unidad institucionalizada y cristalizada.

En su análisis contrastivo de la fraseología en italiano y español, Quiroga (2006) afirma que «la presencia de un elemento léxico común en las UUFF que se ha de enseñar es una ayuda para el aprendizaje por parte del estudiante»[137], aunque añade posteriormente que dicho criterio no puede ser el único en la elaboración de un corpus fraseológico. Es importante que las unidades aparezcan en un contexto real a través de fragmentos de textos como obras literarias, textos periodísticos, etc. El presente trabajo comparte dicha postura en cuanto se considera

[137] Por otra parte, más adelante en su análisis, la autora (2006) defiende que, en lugar de la agrupación por palabras claves, resulta ser más adecuado ordenar las UF en torno a un significado fraseológico parecido.

sumamente insuficiente el aprendizaje mnemónico de las unidades fuera de un contexto real.

Aunque desde un prisma diferente, el posicionamiento de Francisco Núñez-Román posee similitudes con el de Quiroga. En su artículo *Enseñar fraseología: consideraciones sobre la fraseodidáctica del español*, el autor profundiza en la necesidad de enseñar fraseología a hablantes nativos y afirma que «en Educación Primaria, en la que es fundamental mejorar y enriquecer la capacidad comunicativa de los alumnos, será de especial utilidad la presentación de las UF agrupadas por temas o por palabras clave» (Núñez-Román, 2011: 161), y añade que dicho proceso facilita la creación de «mapas conceptuales en los que se presenten las UF que giren en torno a un tema determinado». Su objetivo es claramente didáctico: utilizar el IF de las palabras claves con el objetivo de crear mapas conceptuales con unidades que comparten un tema específico.

Agnieszka Szyndler (2015) considera que los criterios de agrupación por palabras no son «muy eficaces y rentables, dado que dan prioridad al significado compositivo y descartan por completo el significado unitario, esto es, global y figurativo de una locución determinada» (2015: 211). De la misma opinión es Ruiz Gurillo (2000: 261), quien también rechaza este criterio porque no ofrece una visión global de las UF sino que las introduce como si fueran unidades segmentadas en lugar de dejar claro que siempre se reproducen de manera unitaria y que son indisolubles. Como se ha visto, la fijación fraseológica es el «carácter supremo de la fraseología» (García Page, 2016: 39) y, por ello agrupar, la fraseología en función de un lexema único puede ser limitativo y contraproducente.

Fernández Prieto va más allá y considera que la asimilación mnemónica de las UF puede verse afectada si se sigue el indicador de agrupación de palabras claves. Hablando específicamente de un ejercicio estructurado por campos semánticos[138], considera que «este ejercicio nos parece bien

[138] Contenido en Abanico de Chamorro Guerrero (1996:178-179).

como repaso y refuerzo, pero nunca como presentación pues aunque hay un elemento unitario - el color - los significados de las unidades no están relacionados y sería muy difícil memorizarlos» (Fernández Prieto, 2004: 354).

Arianna Alessandro también comparte esta opinión, según la cual enumerar UF en función de uno de sus componentes es muy limitado, puesto que con este criterio no se resaltan las «relaciones semánticas, ni pragmáticas entre las secuencias presentadas, con estos métodos no se ofrecen pistas para su interpretación ni empleo; esto implica que no favorecen una memorización, una asimilación real» (Alessandro, 2011: 227).

Cabe considerar que dicho IF puede ser sustancialmente limitativo y contraproducente si no se completa con otros indicadores y si no se presenta a través de un contexto real. Utilizando las palabras de Penadés: «falta por comprobar la validez de este recurso - y más cuando se apoya en una única palabra de la unidad fraseológica - en la enseñanza de las unidades fraseológicas, frente a las asociaciones que se pueden crear en función de las relaciones semánticas entre unidades fraseológicas» (1999: 28).

4.2.2. Campos semánticos

En la década de 1930, la escuela de autores alemanes y suizos, entre los que destaca Jost Trier, elaboran la teoría de los campos semánticos. Trier (1931) define el campo semántico como un conjunto de lexemas relacionados recíprocamente entre ellos en función de su semántica. En este sentido, se trata de un grupo de palabras relacionadas entre ellas por su significado que, además, comparten la misma categoría gramatical. Por ejemplo, la ira, el trabajo, la muerte, los colores, los insultos, etc.

Las nociones de Trier fueron ampliadas en las décadas sucesivas[139] con el objetivo de demostrar la existencia de una subestructura en un conjunto de vocablos que poseen una misma base significativa. Las palabras no son, en este sentido, unidades aisladas yuxtapuestas, sino un repertorio de unidades organizadas en torno a núcleos establecidos por los hablantes de cada lengua.

Baralo considera que el objetivo principal del proceso de enseñanza-aprendizaje del léxico se basa en «la creación de redes, formales y semánticas, entre las palabras» (Baralo, 2006: 9). Dichas redes, que se constituyen en niveles cognitivos dentro del lexicón de los hablantes, representan una motivación añadida para el aprendizaje del léxico.

> Todo el valor comunicativo del lenguaje, en todas las lenguas, se basa en la capacidad de asociar formas con funciones y significados para darles un sentido. Por ello, la semántica se esparce por todos los niveles lingüísticos, desde el fonológico al pragmático. (Baralo, 2006: 1)

Desde este punto de vista, el criterio semántico debería reinar en los manuales de lengua extranjera. Penadés Martínez define las UF como «combinaciones de palabras que muestran un alto grado de fijación en su forma y en su significado» (Penadés Martínez, 1999: 11), es decir, que al ser unidades estables morfosintácticamente (fijación en la forma), lo son también en su semántica (fijación en el significado). Por ello, el concepto de campos semánticos puede ser trasladado a la fraseología conformando los que se podrían definir como campos semánticos fraseológicos.

En las últimas décadas, se han proporcionado numerosas aportaciones para clasificar el universo fraseológico a partir de criterios semánticos. En un artículo titulado *La descarnada/la secca en la fraseología española e*

[139] Sus representantes europeos más destacados son Greimas, Pottier y Coseriu.

italiana, Vanda Durante (2014) analiza la fraseología que los hablantes italianos y españoles emplean para exorcizar el miedo a la muerte y propone la compilación de un corpus, compuesto por locuciones, paremias y fórmulas, y basado en «microcampos semánticos bilingües» (Durante, 2014: 241).

En dicho campo semántico de la muerte, Durante individua once microcampos semánticos en los que organiza su corpus bilingüe español-italiano. A saber, 1. Agonía: *tener las horas contadas* o *los días contados* ≈ *avere le ore contate* o *i giorni contati*; 2. Descomposición: *dar de comer a los gusanos* ≈ *dare da mangiare* o *fare da cena ai vermi*; 3. El cese: *doblar la servilleta* ≈ *staccare la spina / piegare il tovagliolo*; 4. El sueño: *dormir el sueño eterno* ≈ *dormire il sonno eterno*; 5. El viaje: *irse a la tumba* o *al hoyo* ≈ *andare alla fossa*; 6. Inmediatez: *caerse redondo* ≈ *cadere a corpo morto* o *morto stecchito*; 7. Lugares: *bajar al sepulcro* ≈ *scendere al sepolcro*; 8. Modos: *morir como un perro* ≈ *morire come un cane*; 9. Religión: *pasar a mejor vida* ≈ *passare a miglior vita*; 10. Signos: *con los pies por delante* o *para (a)delalante* ≈ *con i piedi davanti* o *in avanti*; 11. Violencia: *morir en la brecha* ≈ *morire* (o *cadere*) *sulla breccia*.

En este sentido, la autora propone «un modelo centrado en la teoría de los campos semánticos con el fin de establecer un sistema connotativo que reduzca, siquiera parcialmente, la complejidad subyacente a la traducción de español a italiano de las expresiones idiomáticas relativas a la muerte y sus postrimerías» (Durante, 2014: 241).

Otra aportación a una clasificación de la fraseología en función de los campos semánticos proviene de Fernández Prieto quien, en un interesante y ya clásico artículo sobre la fraseodidáctica, insiste en que «la mayoría de los autores consultados también inciden en la necesidad de desarrollar clasificaciones desde el aspecto semántico» (Fernández Prieto, 2004: 353) y que el desarrollo de estas agrupaciones es fundamental para crear ejercicios que aborden la fraseología. De hecho,

la elaboración de los manuales corrobora dicha afirmación, puesto que, como se verá a continuación, existe una tendencia a estructurar y elaborar ejercicios fraseológicos también en función de los campos semánticos.

Penadés Martínez presenta una serie de ejercicios para que el alumnado pueda aprender y memorizar fraseología a partir de las relaciones semánticas que presentan las UF en su manual *La enseñanza de las unidades fraseológicas* (1999). La autora sugiere que su memorización se ve facilitada si se analiza un conjunto de unidades a través de tres tipos de relaciones semánticas: sinonimia, hiperonimia-hiponimia y antonimia. Por ello, cinco de los nueve ejercicios que propone se estructuran en función de los vínculos semánticos que entrelazan entre las distintas unidades fraseológicas (Penadés Martínez, 1999, 55-57).

En este sentido, las unidades que comparten un tipo de relación semántica deberían, según la autora, ser enseñadas a la vez: *llorar como una magdalena*, *llorar a moco tendido* y *llorar a lágrima viva* son hipónimas del hiperónimo *llorar*; en cambio, se rigen por relación semánticas de sinonimia las UF *no valer un cuerno, no valer un pico y no valer un pito*. Finalmente, la antonimia caracteriza las unidades *estar por debajo de* y *estar por encima de,* puesto que las dos UF presentan significado opuesto (Penadés Martínez, 1999: 60 y ss.).

Cabe destacar que dicho IF basado en los campos semánticos, a pesar de ser valioso para impartir fraseología, puede fomentar cierta confusión. Cada unidad posee una carga metaligüística propia y el emisor tiene que tener en consideración otros factores como la situación comunicativa o el registro en el que se emplea una UF u otra (Penadés Martínez, 2015). Por ello, se considera que el IF debe ser combinado con otros, ya que la didáctica de la fraseología no se puede limitar a la mera memorización de unidades en función de su relación semántica.

4.2.3. Unidades didácticas

A continuación, se examinan algunas propuestas para compilar una colección de UF con fines didácticos en función de las unidades didácticas o temas didácticos, lo que está estrechamente ligado al criterio del campo semántico. La gran mayoría de los métodos y los manuales que se emplean en la actual didáctica de una LE se estructuran en función de unas unidades didácticas (Fernández Prieto, 2004: 354) y del nivel del MCER al que se refieren.

A pesar de que la elección de las unidades didácticas pueda variar en función del manual, en el nivel A1, por ejemplo, es muy frecuente encontrar unidades didácticas centradas en el tiempo libre, la familia, la compra, etc. Asimismo, en el nivel A2 se suelen tratar temas como los viajes, el *Made in Italy*, el entretenimiento televisivo, etc. Alrededor de dichas unidades didácticas, y su relación con la gramática, los manuales desarrollan sus contenidos gramaticales, las funciones comunicativas, el léxico, etc.

La sección 6.4.7.3. del MCER se propone seleccionar el léxico en función de áreas temáticas, estrategia didáctica que puede ser trasladada a la enseñanza de la fraseología a través de dos enfoques distintos: léxico y semántico. El primero relaciona la unidad didáctica con la unidad léxica al que pertenece el constituyente léxico principal que compone la unidad: la fraseología se agrupa en función de uno de sus componentes léxicos en lugar de priorizar el significado holístico de la unidad[140].

En este sentido, en la unidad didáctica «la casa», tienen cabida UF italianas como *A casa di Dio*, para indicar un lugar muy lejano y difícil de alcanzar del que se encuentra el hablante; *Andare a letto con le galline*, ir a la cama muy pronto; o *Avere poco sale in zucca,* para identificar a

[140] Este procedimiento es considerado como una alternativa a la agrupación por significado unitario de las unidades y es avalado por Mario García-Page en *VI Congresso Internazionale di Fraseologia e Paremiologia Phrasis Fraseologia e paremiologia: Modelli, strumenti e prospettive* (2019, actas en prensa).

una persona que carece de inteligencia o raciocinio, etc. Los elementos léxicos que se encuentran en las unidades (la casa, la cama, la sal y la calabaza) representan elementos unificadores alrededor de los cuales recopilar fraseología, puesto que se pueden encontrar en una unidad didáctica que trate la casa.

El segundo enfoque, el semántico, pone en relación las unidades didácticas de los manuales con el significado holístico o global que comparten diferentes unidades fraseológicas. En este sentido, alrededor de la unidad didáctica del «trabajo» (solitamente propia de los niveles B1 y B2 del MCER), se podrían incluir UF como *Lavorare sodo*, trabajr duro; *Fare carriera,* conseguir mejorar el recorrido profesional de una persona; *Essere al verde*, no tener dinero; o la cita *Il lavoro nobilita l'uomo*, el trabajo ennoblece. Ambos enfoques (léxico y semántico) aplicados a las unidades didácticas tienen razón de ser y su validez en función del objetivo con el cual se construye el listado de UF con fines fraseodidácticos.

Según Núñez-Román, las relaciones cognitivas entre UF son «un instrumento fundamental para enriquecer el vocabulario y ampliar la organización semántica de los alumnos» (Núñez-Román, 2015: 160-161). El autor considera que se trata de uno de los indicadores más productivos a la hora de impartir fraseología, tanto en fases iniciales, como avanzadas y, aunque el profesorado se refiera a este criterio con un enfoque para la L1, considera que podría llevarse a cabo también en un contexto de lengua extranjera.

Basándose en las teorías didácticas de Baralo (2006) sobre las redes semánticas como construcciones cognitivas que se deben crear para enseñar una lengua, Agnieszka Szyndler critica el modelo de explotación didáctica que pretende agrupar las UF en función de una palabra clave y propone sustituir dichos ejercicios «por otras actividades en las cuales

las UF están clasificadas, o bien según el campo conceptual al que se refieren, o bien por temas» (2015: 201).

Por su parte, Fernández Prieto (2004: 354) argumenta que los ejercicios elaborados en función de elementos semánticos pueden ser de utilidad solo como «repaso y refuerzo, pero nunca como presentación pues aunque hay un elemento unitario -el color- los significados de las unidades no están relacionados y sería muy difícil memorizarlos». En cambio, propone seguir unos criterios que considera más rentables y que denomina agrupaciones por «campos conceptuales». Dicho IF permite diseñar actividades alrededor de temas como la del ocio, el trabajo, la familia, etc., campos conceptuales en torno a las unidades didácticas que suelen presentar su contenido gramatical.

En conclusión, cabe reflexionar sobre si el criterio que se aplica en los manuales para aprender el léxico también se podría considerar en la enseñanza de la fraseología. La propuesta de reunir la fraseología en función de dicho indicador fraseodidáctico aplicado al constituyente léxico principal de la UF no parece acertada dado el carácter composicional de la fraseología. En cambio, se considera más productivo para fines didácticos emplear el significado global de las UF frente al unitario. Desde una perspectiva didáctica, el significado composicional de la fraseología debe primar sobre el significado de cada uno de sus constituyentes puesto que el proceso de memorización se podría ver afectado negativamente.

4.2.4. Uso contrastivo

El parecido formal y, como se ha visto anteriormente, semántico de una palabra entre la LO y la LM representan dos de los factores principales que permiten su descifración y memorización facilitando el proceso de reproducción (Ettinger, 2008). En este sentido, las ventajas de aplicar el uso contrastivo entre UF de dos idiomas como indicador fraseodidáctico han sido abordadas en numerosos trabajos (entre otros, Gómez Molina,

2000; Navarro, 2003; Molina García, 2006; Quiroga, 2006; Alessandro, 2011).

A partir de los universales semánticos[141] (Wierzbicka, 1972, 1992, 1996, 1997) la comunidad académica ha elaborado y estudiado el concepto de los universales fraseológicos[142] que Penadés (1999) define como aquellas unidades transculturales que en lenguas distintas muestran poseer un patrón cultural homólogo. Según la autora, ciertas UF, que expresan experiencias universales o estados de ánimo como el tiempo, la felicidad, la juventud, etc., se memorizan con más facilidad, puesto que «se expresan en una lengua a través de unidades fraseológicas que tienen una correspondencia en otras lenguas por pertenecer a las nociones significadas de un fondo común a diversas culturas» (Penadés, 1999: 37).

En relación a los universales fraseológicos, los estudios contrastivos entre italiano y español «han tenido, hasta el momento, un desarrollo bastante limitado» (Alessandro, 2011: 232), por lo que no se puede afirmar con rigor científico que las similitudes evidentes entre dichos idiomas puedan favorecer a los hispanohablantes a la hora de aprender fraseología italiana. Con todo, se han desarrollado acercamientos (Molina García, 2006 y Alessandro, 2011) en los que se defiende que las similitudes entre idiomas afines deben ser aprovechadas a la hora de enseñar una lengua extranjera, incluyendo no solo las UF que padecen una equivalencia total, sino también las unidades que presentan un grado de equivalencia parcial[143].

[141] Se trata de un metalenguaje semántico universal compuesto por primitivos semántico" (Wierzbicka, 1972), basado en los significados de palabras que puede ser empleado para definir cualquier unidad léxica en cualquier lengua.

[142] Véase Wotjak (1983); Dobrovol'skij (1988, 1998a, 1998b); Luque Durán y Manjón (1998); Gibbs (1999); Dobrovol'skij y Piirainen (2000); Forment Mar (2000); Luque Durán (2004), García-Page Sanchez (2017), María Navarro Paco (2018), Serradilla Castaño (2019) etc.

[143] Con respecto al concepto de equivalencia véase: Gladrow 1993, Dobrovol'skij 2000, 2014; Mellado Blanco 2015 y Corpas Pastor 2000; 2003; 2007.

Estudiar una LE que comparta con la L1 rasgos gramaticales, semánticos o fonéticos ofrece ciertas ventajas debidas a la interlengua y las estrategias de exploración interlingüísticas (Calvi, 2004). La interlengua es uno de los rasgos principales que respaldan el empleo de dicho IF basado en el uso contrastivo entre dos idiomas para impartir fraseología en el aula. Simone (2011) hace suyo dicho IF y añade que:

> en una hipotética escala de dificultad, sería aconsejable empezar con aquellas UF que cuentan con expresiones idénticas disponibles en la lengua materna del aprendiz, las afinidades entre italiano y español posibilitarían un aprendizaje más rápido, puesto que el alumno ya posee cierto conocimiento del significado, de la función y del contexto de empleo adecuado de la secuencia. (Simone, 2011: 234)

La autora propone emplear dicho IF de manera proporcional: cuanto mayor sea el grado de equivalencia entre dos unidades (L1 y LE), menor será el nivel del MCER en el que se deberá de impartir la UF en cuestión.

Siempre desde un punto de vista contrastivo, Juan Ramón Gómez Molina (2000) elabora un trabajo interesante sobre la enseñanza de locuciones en la clase de ELE. Parte de la hipótesis de que las locuciones equivalentes o idénticas entre la L1 y la LE son más fáciles de aprender, puesto que el alumnado dispone de cierto conocimiento de la significación idiomática, de la función expresiva y del uso sociolingüístico de dichas locuciones.

Los resultados de su análisis corroboran, en parte, dicha hipótesis: si por un lado, el estudio contrastivo es positivo para el aprendizaje de los criterios sintácticos y pragmáticos, no es tan prolífico en cuanto a los aspectos semánticos. En este sentido, existe el riesgo de que las unidades que no estén completamente consolidadas en el lexicón mental del alumnado sean incorporadas de manera literal. Como se verá en los siguientes apartados, dichas conclusiones llevarán a Gómez Molina

(2000) a proponer una metodología basada en IF como las relaciones asociativas funcionales, conceptuales y contrastivas.

Otra autora que analiza y defiende el potencial de los estudios contrastivos aplicados a la didáctica de la fraseología es Fasla (1996). Sin embargo, no se refiere a la contrastividad semántica entre dos UF, sino más bien a la morfosintáctica. Según Fasla, dicha técnica permite al alumnado establecer autónomamente la presencia de paralelismos sintácticos entre la L1 y la lengua estudiada para comprobar similitudes o diferencias léxicas entre las UF. La única *conditio sine qua non* para poner en práctica dicho indicador fraseodidáctico de la contrastividad es el buen conocimiento de la lengua nativa por parte del alumnado. Además, la autora considera que la traducción es una técnica eficaz para poner en práctica dicho IF y, teniendo en cuenta la didáctica y la estructura interna de las UF (que denomina modismos), presenta una clasificación basada en niveles lingüísticos en los que distribuye cuatro tipologías de unidades.

El morfosintáctico es el primer nivel y en él se encuentran: a) las unidades variables en persona, género y número, que se formulan preferiblemente en un enunciado afirmativo (*Todo le sale a pedir de boca*), y b) las unidades invariables, que no admiten ningún cambio morfológico o sintáctico (*Del dicho al hecho hay gran trecho*). El segundo nivel es el léxico-semántico y este incluye: a) unidades que admiten tanto la forma afirmativa como la negativa, además de admitir variaciones gramaticales y lexicales (*No estoy muy ducho en esta materia / el traductor está muy ducho en francés*), y b) unidades que se construyen normalmente con una partícula negativa, o que indican la negación de forma implícita (*No es nada del otro jueves / da igual*). Según la autora, este tipo de unidades constituye una minoría y debe ser estudiado en los niveles más especializados del MCER.

Finalmente, cabe destacar la limitación que conllevan los falsos amigos fraseológicos que, juntos con los falsos amigos léxicos, constituyen un problema tanto traductológico como didáctico. En este sentido, puede ocurrir que el alumnado dé por sentado el significado de una UF dado que, en su idioma materno, dispone de otra unidad cuya estructura morfosintáctica es muy similar, pero su significado no coincide completamente o, en ocasiones, puede reflejar imágenes distintas. Julia Sevilla Muñoz y María Sardelli exponen algunos ejemplos en El Trujuman[144], la revista de traducción del Centro Virtual Cervantes: la locución española *ir de punto en blanco* y la italiana *di punto in bianco* son falsos amigos fraseológicos en cuanto la UF española es utilizada para indicar que alguien va vestido muy bien, mientras que la italiana es una locución adverbial traducible en español con *de repente*.

Además, se debe añadir al análisis crítico otra limitación potencial ya que se presenta un alumnado como lingüísticamente uniforme, es decir, que comparten una misma lengua materna. Desde este punto de vista, se deben considerar dos posibles escenarios: por un lado, las clases de ILE que se desarrollan fuera de Italia y, por otra parte, las que tienen lugar en el país.

Salvo excepciones menos representativas, donde el número de extranjeros residentes es muy elevado (véase Luxemburgo[145]), el alumnado del primer caso suele ser lingüísticamente homogéneo, de modo que en las clases de ILE en España la mayoría de los estudiantes suele ser hispanohablante, en Francia francófona, etc. En cambio, en las clases de ILE que tienen lugar dentro del país se encuentran dos realidades que complican el análisis de la realidad: o bien puede ocurrir

[144] El Trujumán (26 de septiembre de 2011) disponible en <https://cvc.cervantes.es/trujaman/anteriores/septiembre_11/26092011.htm>.

[145] Según el Ayuntamiento de Luxemburgo, con fecha 31 de diciembre de 2019, el 70,59% de los residentes en la ciudad son extranjeros (86.360 personas). Más info en: <https://www.vdl.lu/sites/default/files/media/document/Etat%20de%20la%20population%202019_0.pdf>.

que el origen y, por ende, las lenguas maternas del alumnado sean muy variadas y diferentes; o bien, como en las clases de italiano LE para migrantes, a esta primera dificultad se añade el problema del analfabetismo funcional y de la gran distancia de los idiomas de origen con el italiano.

En conclusión, el IF del uso contrastivo ha sido ampliamente tratado en la didáctica de la fraseología y se ha demostrado su validez para estructurar un corpus de unidades fraseológicas con intención didáctica. Asimismo, cabe destacar que presenta limitaciones prácticas relacionadas, por un lado, con los falsos amigos semánticos y, por otra parte, con la no universalidad de las estrategias didácticas, lo que conlleva la no replicabilidad en distintos idiomas. Dicha aproximación se vuelve efectiva solo si el análisis contrastivo se ciñe a un número limitado de idiomas o a un colectivo de estudiantes lingüísticamente homogéneo.

4.2.5. Funciones comunicativas

Jakobson (1975) define las funciones comunicativas como las tareas o roles que los enunciados cumplen en el proceso comunicativo y, por ello, la función primordial de una lengua es consentir una comunicación eficaz entre emisor y receptor. Desde este punto de vista según el cual la lengua es un sistema complejo cuyo fin es la comunicación, otro indicador fraseodidáctico que se debe tener en cuenta es la agrupación de la fraseología estructurada por funciones comunicativas.

A través del léxico simple o las unidades fraseológicas, una determinada comunidad de hablantes utiliza ciertos actos comunicativos para expresar funciones comunicativas como el miedo, la rabia, las molestias, la felicidad o la indiferencia, por ejemplo. Esta perspectiva práctica y enfocada más al uso real de la lengua que a la idea del hablante ideal y, por ello, el presente IF, hace referencia tanto a las relaciones existentes entre los signos lingüísticos, como a las relaciones pragmáticas que se dan

entre los interlocutores y el contexto y en el mismo sistema de la lengua (Bachman, 1990).

La sección 2.1.2. del MCER aborda la competencia comunicativa como la unión de tres subcompetencias: las competencias lingüísticas, las sociolingüísticas y las pragmáticas. Las primeras se realizan en el plano del sistema, e incluyen conocimientos y destrezas léxicas, fonológicas y sintácticas. Las competencias sociolingüísticas, en cambio, se refieren a las condiciones sociales y socioculturales del uso de una lengua que afectan a la comunicación entre representantes de culturas diferentes. Finalmente, las competencias pragmáticas se centran en el «uso funcional de los recursos lingüísticos (producción de funciones de lengua, de actos de habla) sobre la base de guiones o escenarios de intercambios comunicativos».[146] En este sentido, la lengua posee un aspecto práctico y uno más pragmático con los que hablantes y, por lo tanto, alumnado de una LE deben que lidiar.

Desde una perspectiva cognitivista[147], Marta Higuera (1997, 2009) crítica el empleo del IF de los campos semánticos, descrito en el apartado 4.2.4. del presente estudio, al presentar la gran limitación de reunir unidades asociadas a diferentes funciones comunicativas lo que, según la autora, comporta un déficit en la memorización. Para obviar a dicho problema la responsable de Unidad de Cursos y Actividades Formativas del Instituto Cervantes propone emplear como indicador fraseodidáctico la función comunicativa que una unidad fraseológica[148] desarrolla en el discurso.

[146] Para más información véase:
<https://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/cap_02.htm#p212>.

[147] Higuera define el cognitivismo como «un enfoque basado en el uso (el estudio del lenguaje no puede separarse de su función cognitiva y comunicativa), que pone énfasis en la lengua real, por eso no deja fuera las construcciones idiomáticas que el generativismo consideraba ajenas a las reglas generales [Langacker, 1987, 57]» (2006: 11).

[148] La autora (2009, 112 y ss.) habla de unidades léxicas que divide en dos grupos: por un lado las que «coinciden con la unidad "palabra", como por ejemplo: coche, sabiduría, comer, inteligente..., etc.» y, por otra parte, «las que constan de más de una palabra, como por ejemplo: por si acaso, cada dos por tres, chica morena..., etc.». Estas últimas, que en este trabajo se han definido como unidades fraseológicas, son divididas por la docente en tres subgrupos: a) Las frases hechas b) Las combinaciones sintagmáticas c) Las expresiones institucionalizadas.

Según Higuera (2009: 125), se trata de un valioso IF para las clases de LE puesto que «estas fórmulas ayudarán a los alumnos a desenvolverse en muchas de las situaciones en las que se verán fuera del aula».

Refiriéndose a las fórmulas rutinarias, Inmaculada Penadés Martínez (2001) propone tener en cuenta las relaciones semánticas entre las UF para establecer relaciones de sinonimia y antonimia. Según la autora, los signos lingüísticos se asocian entre ellos al estar ligados entre sí por «diversos grados de similitud y fluctúan entre el grado de equivalencia de los sinónimos y el núcleo común de los antónimos» (Penadés Martínez, 2001: 42). La docente sugiere explotar las relaciones de antonimia, hiperonimia-hiponimia y sinonimia entre UF. Asociando los signos lingüísticos que producen estas relaciones, el alumnado podrá elaborar de manera natural «series mnemotécnicas virtuales que pueden ser de gran utilidad para memorizar y, al final, aprender unidades lingüísticas con rasgos especiales como la fijación formal y semántica que caracteriza a las unidades fraseológicas» (1999: 43). En este sentido, la unidad *importar un bledo* puede considerarse sinónima de *importar un comino* y de *importar un pimiento* (Penadés, 1999) y, puesto que todas expresan la misma función comunicativa para expresar indiferencia, la autora propone que se impartan a la vez. A través de este IF elaborado a partir de los vínculos semánticos y comunicativos entre unidades, la profesora elabora una colección de UF a partir de dos diccionarios[149].

En otro interesante trabajo sobre didáctica del español y competencias pragmática, Ana María Ruiz Martínez (2005) presenta un breve listado fraseológico de acuerdo con los descriptores propuestos por el MCER para los niveles A2 y B1. Observando que, como se ha visto en el capítulo precedente de este estudio, el MCER no hace referencia a la fraseología en los niveles más básicos, la autora analiza los descriptores del Marco Común Europeo que pretenden ilustrar diferentes aspectos cualitativos

[149] *Diccionario fraseológico* (Varela y Kubarth, 1994) y *Diccionario para la enseñanza de la lengua española* (Moreno Fernández, 1995).

relacionados con el uso de la lengua y se da cuenta de que hay numerosas funciones comunicativas, como por ejemplo saludar, que obligan al alumnado de los niveles más básicos a «utilizar determinadas unidades fraseológicas (buenos días, ¿qué tal?, ¿cómo andamos?, entre otras)» (Ruiz Martínez, 2005: 569).

Por ello, elabora un listado de UF del ELE para registrar la fraseología que, en función de su significado comunicativo-pragmático, se pueda incluir en una programación didáctica respetando los niveles del MCER[150]. Ruiz Martínez examina entonces doce descriptores ilustrativos del MCER establece una «conexión entre los actos comunicativos del Repertorio[151] para el nivel elemental e intermedio y los fines funcionales que el Marco agrupa en los niveles A2 y Bl» (Ruiz Martínez, 2005: 571). De esta manera, obtiene un corpus básico de UF en función del MCER conformado por 17 fines funcionales para el nivel A2 y 17 para el nivel B1[152]. Como resultado, la autora obtiene un pequeño pero valioso instrumento fraseodidáctico elaborado a partir de un IF específico y que puede, indudablemente, ser de gran ayuda para que profesorado de ELE establezca qué unidades fraseológicas incluir en sus clases en función de los niveles A2 y B1 del MCER.

Como se ha observado, dicho IF permite agrupar el universo fraseológico según las funciones comunicativas que las UF desempeñan en el habla y es bien recibido por la academia. Finalmente, se puede afirmar que, de acuerdo con Agnieszka Szyndler (2015: 212), «resulta imposible introducir las UF en clase sin recurrir a las funciones comunicativas que pueden desempeñar en el discurso» puesto que todo tipo de actividad fraseodidáctica tiene que emplear el principio

[150] En este trabajo la autora se centra solo en los niveles A2 y B1.

[151] En particular, la autora menciona que ha empleado como guía el Repertorio de funciones comunicativas del español, de Gelabert, Martinell, Herrera y Martinell (1996).

[152] Para observar el corpus véase Ruiz Martínez (2005: 572-579).

comunicativo/pragmático siendo imprescindible para el alumnado conocer las funciones discursivas que cada UF desarrolla.

4.2.6. Idiomaticidad y equivalencia

Como se mencionó a lo largo del capítulo II de la presente tesis doctoral, la relación entre el significado literal y el composicional de una unidad fraseológica es, a menudo, una de las características principales de los fraseologismos y dichas unidades idiomáticas «ya no se ven como excepciones o irregularidades, sino como construcciones propias que pueden compartir ciertos aspectos con expresiones plenamente productivas (y viceversa)[153]» (Corpas Pastor, 2022: 3).

Desde una perspectiva didáctica, el concepto de equivalencia está estrechamente ligado con la idiomaticidad de la fraseología puesto que juega un papel fundamental en la memorización de ciertas unidades. Como se verá a lo largo de este apartado, el criterio de equivalencia y el de idiomaticidad en la didáctica de una lengua extranjera se deben considerar como un *unicum* en cuanto proporcionales. Cuanto menos idiomática sea una unidad fraseológica, más transparente será su significado y el alumnado identificará rápidamente un equivalente en su idioma. En cambio, si la composicionalidad o la idiomaticidad de la UF no permite establecer un equivalente inmediato, será deber del profesorado hacer transparente la unidad para que el alumnado pueda descifrar su significado.

Por ello, idiomaticidad y equivalencia se tratan en este trabajo como un indicador fraseodiáctico único, ya que en el aprendizaje de una lengua extranjera son principios que ocurren siempre en combinación: por cada unidad idiomática aprendida, el alumnado busca en seguida su equivalente en la lengua de orígen con el fin de facilitar su memorización.

[153] Traducción propia: In this constructional network, traditional idioms are no longer viewed as exceptions or irregularities, but as constructions of their own that may share certain aspects with fully productiveexpressions (and vice versa).

La idiomaticidad fraseológica, propiedad no necesaria de las UF, hace que el significado de la unidad no sea deducible por sus componentes sino solo en su conjunto. Bosque y Gutiérrez-Rexach explican dicho fenómeno mediante lo que denominan «principio de composicionalidad»:

> En casos como *tomar el pelo* o *meter la pata* tenemos grupos sintácticos a efectos gramaticales, pero a efectos semánticos no podemos analizar composicionalmente su significado, es decir, no podemos combinar la denotación de *tomar*, la de el y la de *pelo* y obtener el significado del conjunto de acuerdo con los principios sintácticos con los que obtenemos el significado de otros sintagmas verbales. (Bosque y Gutiérrez-Rexach, 2009:114)

Por otra parte, el concepto de equivalencia resulta un fenómeno complejo que debe ser examinado con detenimiento. Mellado Blanco (2015: 154-156) teoriza la existencia de tres planos distintos en los que este fenómeno actúa: sistémico, que no depende de un contexto determinado ni de la deixis de tiempo y espacio; textual, que depende del contexto en el que se inserte y su equivalente solo puede encontrarse en función de un texto específico; y, por último, lexicográfico, que se sitúa en un plano intermedio entre el sistemático y el textual.

Corpas Pastor ha teorizado dicho aspecto en varios de sus artículos sobre la problemática de la traducción de la fraseología[154] y señala que el proceso se ejemplifica por medio de tres fases consecutivas: identificación, interpretación y búsqueda de correspondencia. Así, la autora describe este proceso como:

> Tras las fases de identificación y posterior interpretación, el traductor procede al establecimiento de correspondencias, primero en el nivel lexicológico, el plano teórico, para, a

[154] Véase, entre otros: Corpas, 2000; así como diversos artículos de esta autora recopilados en Corpas, 2003.

> continuación, bajar a los niveles textual y discursivo, en una palabra, al plano real. (Corpas Pastor, 2003: 216)

En el presente trabajo doctoral se considera que los procesos que la autora define y que atañen al proceso traductológico se dan de la misma manera en el proceso de aprendizaje de una lengua extranjera. Si bien de manera inconsciente, el alumnado de una LE utiliza las mismas técnicas para encontrar equivalentes fraseológicos en su idioma materno y así poder fijar la unidad en su propio lexicón mental.

En este sentido, en el primer nivel lexicológico, el traductor/alumnado de LE tendrá que lidiar con los grados de equivalencia. Corpas (2003: 206-208) identifica tres grados de equivalencias[155]: total, parcial, nula y aparente. La *equivalencia total* se da cuando se crea un traslema[156], es decir, cuando todos los términos que componen una unidad se pueden sustituir de manera literal sin alterar el significado unitario o idiomático de la unidad. En este sentido, se supone que la motivación de las UF que presentan equivalencia total es la misma, dado que disponen de orígenes comunes o que han sufrido procesos de préstamos o calcos lingüísticos.

Por otra parte, la *equivalencia parcial* se da cuando, en palabras de la autora:

> Una locución de la LO se traduce por una locución en la LM que presenta diferencias de significado denotativo o connotativo, o que pertenece a una variedad (diastrática, diafásica y diatópica) distinta; o es traducida por una unidad léxica simple carente de

[155] Cabe destacar que, como apunta Timofeeva (2012), ya Solodukho (1977) y Raikhshtein (1980) habían teorizado una escala gradual de equivalencias.

[156] En este caso, la autora emplea la terminología de Santoyo y Rabadán, (1991) quien define un traslema como una «unidad bitextual de cualquier tipo o nivel constituida por un mismo contenido y dos manifestaciones formales diferenciadas pero solidarias y cuya existencia depende de la relación global de equivalencia subyacente a cada binomio textual T[exto]O[rigen]-T[exto]M[eta]» (Santoyo y Rabadán, 1991: 322, en Corpas, 2003: 205).

> los valores expresivos de la locución en cuestión. (Corpas, 2003: 207)

En este sentido, la UF traducida y la original se pueden considerar homólogas desde un punto de vista comunicacional, aunque será el contexto a establecer si se ha dado un caso de infratraducción o sobretraducción, es decir, la ausencia o la exageración de aspectos semánticos, estilísticos u otros presentes en el texto meta. Dichas incoherencias respecto al texto original se deben, según la autora, a la ausencia de correspondencias entre las imágenes subyacentes o los significados denotativos o connotativos.

Con respecto a la *equivalencia nula*, Corpas Pastor (2003) afirma que esta se da cuando ciertas UF no disponen de equivalentes en la lengua meta. En dichos casos, el traductor/alumnado de una LE se ve obligado a recurrir a una paráfrasis o a otras técnicas que no conservan la denotación, pero sí la connotación original. Finalmente, la *equivalencia aparente* se refiere a «equivalencias que se dan cuando los elementos constitutivos de las unidades del fraseologismo en la lengua origen y lengua meta presentan similitud formal, pero diferencia semántica» (Corpas Pastor, 2003: 208-209), es decir, se corresponde con la relación de equivalencia que presentan los que se han denominado falsos amigos fraseológicos o unidades fraseológicas pseudoequivalentes (Trivić, 2011).

Marta Higuera (1997) reflexiona sobre el criterio de idiomaticidad y de equivalencia con un enfoque directo a la fraseodidáctica. Con respecto al proceso didáctico, la autora opina que se deben considerar solo las unidades que no son equivalentes entre la lengua de origen y la lengua extranjera, puesto que «las que son propias de cada lengua, crean muchos problemas y son las que fundamentalmente debemos enseñar» (Higuera, 1997: 47). En cambio, las unidades que son universales, es decir las que derivan del conocimiento del mundo, así como las no idiomáticas

no es preciso enseñarlas puesto que suelen disponer de un alto grado de equivalencia.

Dicha observación se comparte solo parcialmente en el presente trabajo y, para profundizar en este debate, se expondrán algunos ejemplos. Si se considera el dominio de la naturaleza y el subdominio de la fauna, en la cultura italiana el oso es considerado como un animal solitario, que vive en las montañas y que puede llegar a ser agresivo. De hecho, la locución italiana *essere un orso* (ser un oso) es utilizada para definir una persona con mal genio, antisocial, que ama estar solo. Sin embargo, en alemán el oso es identificado con la fuerza y la temeridad, por lo que la locución alemana correspondiente a ser un oso tiene un significado completamente diferente, convirtiendo la UF en un falso amigo fraseológico. En este sentido, el léxico que una determinada comunidad lingüística emplea y, por ende, su fraseología, está estrechamente relacionado con la manera de percibir la realidad y esta puede ser muy diferente entre un idioma u otro.

Saracho Arnáiz (2015) trata la idiomaticidad como una herramienta fraseodidáctica para desarrollar la competencia fraseológica en la clase de ELE, y propone establecer unos grados de dificultad de las UF en función de la idiomaticidad: las unidades más opacas son más complicadas de descifrar para el alumnado y, por ello, habrá necesariamente que enseñarlas en los niveles más avanzados. En palabras de autora: «cuanto más opaca es una unidad, más difícil es llegar a su significado a través de la forma. Por el contrario, una unidad más transparente hace que sea fácilmente interpretable a partir de su estructura externa» (Saracho Arnáiz, 2015: 927).

La autora considera que se deben adoptar criterios que «no dificulten en exceso el aprendizaje de dichas unidades, ya que el sentido figurado de ciertas unidades puede llegar a ser totalmente opaco» (2015: 927) en los niveles iniciales. En este caso, Saracho Arnáiz propone seleccionar

unidades semitransparentes que, en su opinión, son las más adecuadas para la enseñanza en los niveles más elementales. Para ello, propone empezar con los somatismos, unidades fraseológicas que contienen por lo menos un lexema referido a partes del cuerpo, ya que suelen ser semitransparentes y, por lo tanto, de fácil aprendizaje para el alumnado, al poseer un significado que es interpretable sin esfuerzo.

No obstante, si se evalúan dos idiomas cercanos como el italiano y el español, los somatismos no son siempre tan transparentes como opina la autora. A pesar de que la similitud léxica entre estas dos lenguas afines sea muy alta, las unidades fraseológicas a menudo pueden ser poco transparentes, o incluso opacas, para un hablante nativo de la otra lengua. En italiano, por ejemplo, las locuciones adverbiales *essere una persona alla mano* y *essere una persona in gamba,* a pesar de ser somatismos en cuanto uno de sus componentes es una parte del cuerpo (*mano* y *gamba*), poseen un significado no descifrable para un nativo hispanohablante, en cuanto no hay equivalencia en la lengua de orígen. En este sentido, *essere una persona alla mano* no tendría en el español peninsular un equivalente directo, y se podría traducir con *ser una persona cercana*; mientras que *essere una persona in gamba* en español se podría interpretar como ser experto, hábil o mañoso haciendo una determinada acción, o incluso, estar en buena salud.

Lo mismo ocurre con la fórmula *In bocca al lupo*, unidad que se utiliza para desear suerte a alguien que se aproxima a realizar una tarea que conlleva cierto esfuerzo y que suele estar sujeta a una voluntad ajena a la propia, como un examen o una entrevista de trabajo. Dicha fórmula expresiva para desear suerte (Corpas Pastor, 1996) es opaca para el alumnado, en cuanto su significado literal es *en la boca del lobo*, es decir, enfrentarse a una muerte segura. Sin embargo, suele ser enseñada ya a partir del nivel A1 del MCER, el inicial y más básico en cuanto su alta frecuencia de uso en la lengua hablada, lo que la convierte en una unidad fundamental para quien se acerque al italiano.

En conclusión, a pesar de las aportaciones que plantean estructurar una colección de UF en función del IF de la equivalencia y/o de la idiomaticidad, se considera que dicho indicador si bien es fundamental, parece estar subordinado a otros factores como la frecuencia (véase *In bocca al lupo*, muy opaca pero muy frecuente y, por lo tanto, enseñada ya a partir de los niveles bajos).

4.2.7. Motivación

El concepto de motivación fraseológica ha sido ampliamente debatido en los estudios de fraseología (Corpas Pastor, 1996; Zuluaga, 1980; García-Page, 2008; Pamies Beltrán, 2014; etc.). Como se ha detallado en apartados anteriores (cap. II), con el término motivación el presente estudio hace referencia a la existencia de una relación diacrónica entre el significado literal y el figurado de una unidad fraseológica (Pamies, 2014), es decir, una característica etimológica de la lengua que no pone en relación el significante con su significado (idiomaticidad) sino dos significados: el figurado y el literal.

Para ejemplificar, la unidad coloquial Me la trae al pairo, muy empleada en la lengua hablada por las generaciones más jóvenes, encuentra su motivación en la terminología marinera. El DLE proporciona la siguiente definición para el verbo pairar: «Dicho de una nave: Estar quieta con las velas tendidas y largas las escotas». Por tanto, se puede suponer que la motivación de dicha unidad viene de la sensación de tranquilidad y espencieradez que la tripulación de un barco experimenta en alta mar cuando, para descansar, suelta las escotas y se despreocupa de su labor diaria.

Sin embargo, las derivaciones diacrónicas no son conocidas y la motivación original es oscura al hablante contemporáneo en algunas

ocasiones. El hecho de que se desconozcan no significa que no existan[157], sino que el significado de uno o más componentes de la misma se ha incrustado en la unidad fraseológica cristalizandose. De hecho, si ciertas locuciones permanecen oscuras al hablante es debido a un conocimiento insuficiente y no a una ausencia efectiva de motivación (Robert Martin, 1996). En otras ocasiones, la motivación fraseológica está estrechamente ligada a la cultura de una determinada comunidad de hablantes por lo que, frente a los universales fraseológicos, ciertas unidades presentan particularidades culturales específicas.

Por otra parte, la motivación de una unidad fraseológica no tiene por que ser antigua. La UF *He venido a hablar de mi libro,* pronunciada por el escritor Francisco Umbral en 1993 en el programa en directo «Queremos saber» de Televisión Española, se ha hecho viral, convirtiéndose en una cita que tiene más de 17 millones de resultados en la búsqueda de Google[158] y se emplea, mayoritariamente, para afirmar que se está promoviendo la autorreferencialidad de un discurso personal[159].

La motivación es, sin duda, un factor que resulta particularmente estimulante para el alumnado y, por lo tanto, para la fraseodidáctica, que es objeto de estudio de este trabajo doctoral. De hecho, «las frases hechas cuentan a su favor con la «curiosidad» que despiertan y con la espontaneidad que imprimen al discurso» (Forment Fernández, 1998: 346) y la etimología y el origen de los vocablos tiene una componente atractiva y estimulante para el alumnado.

Dicha motivación no es arbitraria. A veces, puede ser descubierta en una suerte de caza al tesoro, lo que puede animar al alumnado a buscar por sí mismo los posibles significados composicionales, antes de encontrar la

[157] Según Pamies Bertrán (2014) los estudios españoles sobre motivación fraseológica no son avanzados y se le ha dedicado escasa atención a diferencia de otros idiomas como el ruso, el francés y alemán (Mokienko, 1980 y 2007; Buridant, 1989 y 2007, entre otros).

[158] Búsqueda realizada el 10 de julio de 2019.

[159] Para más ejemplos véase el análisis de Pamies Bertrán (2014), quien analiza la motivación de varias unidades fraseológicas, citas y locuciones.

respuesta en el profesorado o en los diccionarios. Como afirma Serradilla Castaño, es posible que «si -un estudiante- sabe que cuando un torero le echa un capote a otro, le está ayudando en una situación comprometida, le será más fácil aplicarla fuera de ese mundo taurino» (2000: 658). Cabe. por tanto, analizar la motivación como un posible indicador fraseodidáctico a seguir para enseñar fraseología en las clases de lengua extranjera.

A pesar de que dicho aspecto de las unidades fraseológicas podría ser un indicador interesante a escoger a la hora de elegir los patrones para elaborar un listado fraseológico específico para la didáctica de la fraseología como LE, la mayoría de los estudios al respecto han demostrado su irrelevancia como se verá a continuación. En este sentido, es preciso detallar las razones por las que la gran parte de la comunidad académica no atribuye funciones didácticas a la motivación, así como los aspectos positivos que dicho IF podría aportar a la elaboración de una colección de UF con fines fraseodidácticos.

La separación metodológica y metalingüística que se da entre el aspecto diacrónico y el sincrónico de una lengua lleva a algunas escuelas a tachar como irrelevantes los datos etimológicos para el análisis y el estudio sincrónico de un idioma. Heine (1997), por ejemplo, razonaba sobre cómo la motivación presenta un carácter subjetivo que varía en función de la perspectiva del hablante en función de un contexto espacial y temporal. En palabras del autor «the synchrony/diachrony distinction derives from the perspective adopted, rather than the facts considered[160]» (Heine, 1997:4)

Casadei (1996) incide, por un lado, en el interés que despierta la motivación en el alumnado aunque, por otra parte, afirma que aunque la motivación semántica de una UF sea cierta y plausible, es

[160] «La distinción sincrónica /diacrónica deriva de la perspectiva adoptada, más que de los hechos considerados». Traducción propia.

indudablemente inútil. Según la autora italiana, la motivación de una unidad fraseológica no aporta nada sobre el uso y el aprendizaje de las unidades fraseológicas, ya que se pierde en la diacronía y, por ende, se vuelve irrelevante también desde un punto de vista sincrónico.

La unidad *ser un lobo con piel de cordero* sigue siendo empleada correctamente por una comunidad de hablantes aunque no se sepa que la frase tiene un origen bíblico, en Mateo 7:15-16[161], y que, aún antes, el fabulista Esopo (VI sec a.C.) había ideado una historia en la que un lobo disfrazado con la piel de una oveja quiso pasar inadvertido entre los corderos, pero su intento fracasó y fue devorado por unos ejemplares de su propia especie. Es más, el hecho de que incluso para los hablantes nativos la motivación de la mayoría de las UF sea oscura corrobora lo dicho.

Según Vanda Durante (2014: 242), el hecho de que el alumnado conozca la motivación de las UF no es relevante, puesto que «a cualquier hablante nativo se le puede escapar la motivación semántica que ha ocasionado el significado composicional de alguna locución». Además, la autora hace hincapié en que los estudios sobre los orígenes diacrónicos (históricos, literarios o mitológicos) han venido considerándose como una «indagación infructuosa» (Durante, 2014: 242).

De acuerdo con estas últimas propuestas, el presente trabajo considera que no es preciso conocer la motivación de unidades como *avere le mani bucate* (ser un/a manirrota) o *avere le braccia/braccine corte* (ser de la cofradía del puño cerrado)[162] para descifrarlas, memorizarlas y emplearlas correctamente en un contexto natural. Por otra parte, es posible que el alumnado tenga más facilidad en retener y memorizar las unidades fraseológicas cuya motivación sea más intuitiva. De modo que, si el docente hace notar que, por ejemplo, se dice *avere le mani bucate*

[161] «Guardaos de los falsos profetas, que vienen a vosotros vestidos de ovejas, pero que por dentro son lobos rapaces. Por sus frutos los conoceréis» (Mateo 7:15-16).

[162] La traducción literal sería *tener las manos con agujeros* y *tener los brazos cortos*.

porque un hueco en las manos hace que el dinero se vaya rápidamente y que *avere le braccia corte* no te permite sacar la cartera del bolsillo, convirtiéndote así en un tacaño, posiblemente su memorización (aunque esto no afecte al uso correcto en contexto), será más efectiva gracias a procesos de asociaciones de ideas.

Por otra parte, la motivación juega un papel relevante en el proceso de aprendizaje de las UF, porque causa un cierto tipo de emociones en el alumnado. Tal como defiende Paolo Balboni, profesor emérito de la Università Ca' Foscari di Venezia y mayor exponente de la Escuela Veneciana de Glotodidáctica[163]:

> Le emozioni sono risposte adattive (in senso darwinistico, cioè finalizzate alla sopravvivenza e all'adattamento) della mente alle pressioni esterne: quindi è fondamentale tener conto delle emozioni, cioè delle reazioni psico-fisiologiche a eventi quali un corso di lingua, la difficoltà a comunicare nella L2 parlata nell'ambiente, la globalizzazione con la sua imperiosa richiesta di inglese lingua franca, l'obbligo di studiare francese o tedesco alla media ecc.).[164] (Balboni, 2013b: 11)

El papel que juegan las emociones en el proceso de aprendizaje de una LE es, según Balboni, fundamental para el desarrollo del mismo aprendizaje. De hecho, en las últimas décadas, se han intensificado las relaciones que pretenden establecer modelos que definan la relación existente entre las emociones y el aprendizaje (entre ellos Cardona, 2001,

[163] Fundada por Giovanni Freddi, la Scuola Veneziana di Glottodidattica se propone superar el concepto de glotodidáctica como mero instrumento de la lingüística aplicada para analizar las repercusiones psicológicas que se realizan en el alumnado y, en particular, las emociones. Para más información véase Balboni (2010, 2011, 2012, 2013a, 2013b).

[164] «Las emociones son respuestas adaptativas (en el sentido darwinista, es decir, dirigidas a la supervivencia y la adaptación) de la mente a las presiones externas: por lo tanto, es esencial tener en cuenta las emociones, es decir, las reacciones psicofisiológicas a acontecimientos como un curso de idiomas, la dificultad de comunicación en la L2 hablada en el entorno, la globalización con su imperiosa demanda de inglés como lengua franca, la obligación de estudiar francés o alemán en la ESO, etc.». Traducción propia.

2012; Abbaticchio, 2002; Borello, 2005; Caon, 2008, 2009, 2010, 2011; etc.)[165].

Por todo ello, y teniendo en cuenta las opiniones negativas previamente esbozadas sobre el uso de dicho IF, el autor de la presente tesis doctoral escribió un correo directamente al profesor Balboni para preguntar su opinión con respecto al empleo de la motivación como indicador fraseodidáctico y el profesor confirmó que, desde su punto de vista, se trataba de un criterio «insignificante».

Cabe destacar igualmente que la etimología suele fascinar y atraer al alumnado que consigue saciar su curiosidad innata de conocer los orígenes históricos, literarios, religiosos o bien mitológicos de una UF. Descubrir el orígen de una UF es estimulante para el alumnado. Como afirma Saracho Arnáiz (2015: 927), «cuando en clase de ELE exponemos a los alumnos las UF idiomáticas, la reacción es siempre de curiosidad y de elevada motivación, elemento clave para el éxito del aprendizaje». No cabe duda de que el proceso de aprendizaje debe ser variado, placentero y estimulante, lo que significa que el profesorado debe seducir al estudiante para que le guste lo que estudia a través de métodos y ejercicios motivadores en las unidades fraseológicas ofrecen un terreno naturalmente fértil, que el profesorado tiene que aprovechar. Con todo, hasta donde se sabe, aún no se ha realizado un estudio empírico para demostrar dicha hipótesis.

Si bien la motivación de las UF ha despertado y sigue despertando interés de las investigaciones lingüísticas y los textos más divulgativos[166], a pesar de que se le reconozca el poder de fomentar la curiosidad del alumnado, se considera que su aplicación a la fraseodidáctica es complementaria y subordinada a otros criterios. Identificar la motivación significa, cuando

[165] En particular, Caon ha trabajado sobre el impacto del placer en el aprendizaje lingüístico empleando como medio vehicular la música (Caon, 2011) y el deporte (Caon y Ongini, 2008).

[166] Entre ellos destaca el interesante compendio etimológico-folcrórico de Néstor Lujan Fernández, *Cuento de cuentos* (1992) en el que se recopila la motivación de numerosas unidades fraseológicas.

es posible, añadir un dato más, una curiosidad que puede ser interesante para la atención del alumnado y, por ende, promover la memorización de la UF (Balboni, 2010, 2011, 2013). Teniendo estas premisas en consideración y siguiendo las indicaciones proporcionadas por la academia, se descarta este indicador para elaborar una metodología propia en el presente trabajo y se limita a considerar dicho IF como subordinado a otros.

4.2.8. Frecuencia

Entre los criterios a analizar en esta investigación doctoral destaca el de la frecuencia fraseológica, puesto que es ser uno de los indicadores más debatidos (Prado Aragonés, 2004; Núñez-Román, 2015) a la hora de elaborar una colección de UF con aplicación fraseodidáctica. Los hablantes nativos no dominan ni la totalidad del léxico, ni de la fraseología de su idioma. Sin embargo, se sienten completamente aptos para comunicar en dicha lengua (FUNDEU, 2011). En este sentido, de todo el léxico que una lengua ofrece, los hablantes solemos utilizar de manera productiva solo un pequeño porcentaje y dicha parte es compartida por un gran número de hablantes del mismo idioma. Este reducto es lo que se ha definido como disponibilidad léxica (López Morales,1973): se conoce de forma pasiva muchos más vocablos y fraseologismos de los que empleamos en la producción activa del habla. La FUNDEU recuerda que:

> De acuerdo a un cálculo de la Academia Española de la Lengua actualizado al 2010, mientras «un ciudadano medio utiliza entre 500 y 1000 palabras» del español para comunicarse cotidianamente, los jóvenes usan un 25%, «algo más de 240». El castellano cuenta con casi 100.000 vocablos, o sea que, de ese

> gran abanico de posibilidades, utilizan un 0,03 % (FUNDEU, 29.12.2011)[167].

De todo el léxico del que dispone un idioma, existe una parte muy reducida que los hablantes utilizan con más frecuencia y que, al ser la más utilizada, viene a ser la que la gran mayoría de los hablantes reconocen. Cuanto mayor sea la frecuencia de una determinada construcción más más fuerte será su arraigo en el trabajo neuronal (Hoffmann, 2013: 315). En otros términos, cuanto más frecuente sea el uso de una combinación de palabras, más alta será la probabilidad de que se cristalice y que los hablantes la reproduzcan tal como es, sin alterar ninguno de sus constituyentes, ni su forma. Con el paso del tiempo, su frecuencia de uso hará que una comunidad de hablantes la reconozca como propia de su patrimonio lingüístico y dicha expresión se convertirá en institucional o convencional en dicha comunidad. Por decirlo en palabras de Corpas Pastor (1996: 22): la frecuencia no es nada más que un fenómeno de repetición diacrónica que desemboca en fijación.

Desde una perspectiva fraseodidáctica, la frecuencia de uso puede ser un IF válido, puesto que aquella parte de la fraseología más frecuente permitiría al alumnado comunicarse de manera productiva con un hablante nativo. La frecuencia es también un factor que se debe tener en cuenta al decidir la progresión con la que se presentará el vocabulario. En el capítulo 6.4.7.3. del MCER se propone «seguir unos principios léxico-estadísticos que seleccionen las palabras más frecuentes en recuentos generales y amplios o las palabras que se utilizan para áreas temáticas delimitadas». Asimismo, autores como Sinclair y Renouf (1985: 154-155) y Nation (2001: 20-21) insisten en que, antes de dedicar esfuerzos a aprender palabras menos frecuentes, el alumnado deberá de adquirir un dominio estable de las formas más empleadas.

[167] Véase: <https://www.fundeu.es/noticia/para-hablar-los-jovenes-utilizan-solo-240-palabras-6835/>.

Como apuntan García Salido, Vincze, Orol González y Alonso Ramos (2016, 39), «uno de los criterios fundamentales para decidir qué ítems del léxico deberían incluirse en un programa de enseñanza de una lengua extranjera/segunda es la frecuencia de uso». Esta hipótesis puede ser también útil para la fraseología, siendo esta última objeto de estudio justamente por su alta frecuencia y recibiendo especial atención en la didáctica de una lengua extranjera. Con todo:

> A pesar de que, debido a su difusión y alta frecuencia de uso en los intercambios comunicativos, las UF se han reconocido como un contenido imprescindible en la enseñanza y aprendizaje de una lengua extranjera, su didáctica en el aula de E/LE e I/LS sigue estando afectada por importantes lagunas. (Alessandro, 2011: 17-18)

En ocasiones, el profesorado de ILE se encuentra con manuales donde el léxico resulta demodé e incluso anacronistico para el italiano contemporáneo. La frecuencia suele ser desatendida en la elaboración de los textos de LE y, de acuerdo con Ruiz Gurillo (1994), las gramáticas deberían de tener más en cuenta la frecuencia en su labor freseográfica. Hasta el momento, la frecuencia de las UF italianas en la lengua escrita y hablada ha sido poco estudiada, y aún no se dispone de un corpus fraseológico ordenado por dicho criterio[168] (Núñez-Román, 2015).

A la hora de seleccionar las posibles unidades destinadas al aprendizaje, Alessandro (2011) considera que la frecuencia de aparición en un corpus debe ser uno de los factores principales. En este sentido, la autora defiende que «el criterio de la frecuencia de uso en corpus puede proporcionar indicaciones útiles a la hora de seleccionar las secuencias para la enseñanza - añadiendo en seguida que, a pesar de cuanto dicho

[168] Cabe destacar a este respecto el trabajo de Domenico Russo (2010): MdD. Modi di dire. Lessico italiano delle collocazioni. En ello el autor ciñe su trabajo a la esfera de las colocaciones y las recoje y organiza (su calsificación se centra en cuatro tipos principales) en función de su alta frecuencia de uso y cohesión semántica.

- no siempre es aplicable precisamente por la escasez de trabajos de este tipo, especialmente por lo que se refiere a la fraseología italiana y española» (Alessandro, 2011: 221).

Por su parte, Martínez Pérez y Plaza Trenado (1992: 124) opinan que la frecuencia es el criterio que debe prevalecer en la enseñanza de un idioma extranjero y, en función de ella, elaboran una serie de actividades con dificultad progresiva pensadas para el alumnado de niveles intermedios-avanzados de ELE. De la misma manera, Francisco Núñez-Román (2015: 160) considera que «uno de los criterios más usados para determinar la importancia de una UF es su frecuencia de uso [...] pero lamentablemente no hay aún ningún repertorio fraseológico organizado según este criterio».

Llegados a este punto, cabe destacar que los estudios sobre disponibilidad léxica empezados por Gougenheim en 1956 y llevados al mundo hispánico por López Morales en 1973 a través de un estudio realizado en San Juan de Puerto Rico[169], demuestran que la frecuencia *per se* es un dato de interés solo si se aplican otros factores a un corpus específico. Aunque algunas expresiones idiomáticas pueden ser poco frecuentes desde el punto de vista estadístico, sí forman parte del lexicón de un hablante nativo, el cual las emplea correctamente desde un punto de vista pragmatismo y las percibe como parte integrante de su lengua.

Todo ello hace que, incluso las menos frecuentes deben ser tenidas en cuenta en el proceso de aprendizaje/enseñanza de una L2 y añade que las UF que se deben incluir en la programación del profesorado son «aquellas que puedan resultar aceptables por un nativo, sin necesidad de que tengan obligatoriamente una alta frecuencia de uso en un corpus en concreto» (Molina García, 2006: 136).

[169] Véase Zuccalà (2013).

Para asignar una UF adecuada en un determinado nivel del MCER es necesario disponer de indicadores fiables. Así, Núñez Cabeza, crítica la excesiva libertad que se le ofrece al profesorado y denuncia «la falta de criterio al seleccionar las expresiones que proponemos los profesores y los manuales» (Núñez Cabeza, 2002: 166). Para demostrar su postura, el autor elabora un estudio en el que examina la frecuencia de más de 336 unidades y las cataloga como alta, normal, o baja, a través de un sistema de evaluación. Sus resultados corroboran la hipótesis, puesto que muchas de las unidades que presentan baja frecuencia aparecen normalmente en los manuales de ELE, con lo que se demostraría una falta de adecuación de los mismos.

Leal Riol (2013) va más allá de la frecuencia y llega a proponer una selección de UF para la didáctica mucho más drástica en la que, además de excluir las unidades que presentan baja frecuencia, propone también eliminar de un supuesto corpus fraseodidáctico «las que expresan valores semánticos muy determinados que resultan difíciles de parafrasear, así como las que contienen palabras diacríticas o nombres propios» (2013: 167).

A pesar de las diferentes propuestas, es evidente que la academia comparte dos ideas firmes: en primer lugar, considera la frecuencia de las UF como un indicador indispensable a la hora de elaborar una colección de UF con fines fraseodidáctico y, en segundo, opina que las unidades con baja frecuencia no deberían conformar este corpus, por lo menos en los niveles básicos e intermedios. En este sentido, el indicador de la frecuencia sigue siendo «uno de los criterios más usados para determinar la importancia de una UF» (Núñez-Román, 2011: 160) y, de acuerdo con Daša Stanič (2014), es necesario profundizar en cada idioma y elaborar listados que tomen y establezcan las unidades más frecuentes, más utilizadas y más útiles para el alumnado de ILS.

Con todo, la frecuencia plantea ciertos escollos que deben ser tenidos en consideración. Penadés Martínez (2015) publica un artículo en el que evalúa algunas de las implicaciones de la frecuencia[170] de las locuciones en la elaboración de un diccionario y pone en evidencia algunas críticas. En su búsqueda de los criterios para la elaboración de un diccionario de locuciones, Penadés Martínez deja patente que la frecuencia fraseológica está sujeta a distintos factores y uno de estos es la dimensión temporal que se escoja. En palabras de la autora: «la frecuencia de las locuciones está condicionada por la concepción que se mantenga del concepto estado de lengua actual» (Penadés Martínez, 2015: 254). En el esfuerzo de definir el concepto de lengua actual, cuyos límites son difícilmente fijables por criterios objetivos, la autora afirma que ciertas locuciones que se emplean todavía hoy en día tienen su origen en los principios del siglo XX y, por ende, es necesario ampliar el espectro temporal hasta dicha época.

La autora insiste en que el cálculo de la frecuencia de las locuciones es un aspecto complejo y reconoce tres factores principales que complican la tarea. En primer lugar, muchas locuciones son de difícil ubicación en una clase, lo que complica su posición y, por consiguiente, su estudio pormenorizado. Por otra parte, y como sucede con las unidades léxicas simples, Penadés Martínez analiza el fenómeno de la polisemia y homonimia en las locuciones y concluye que no basta con elaborar el recuento de una cadena de significantes en un corpus a través de en una base de datos electrónica. Unidades como *meter caña* puede llegar a tener hasta seis significados distintos, a saber y siguiendo a la autora (Penadés Martínez, 2015: 261-262): «Golpear», «Atacar criticando», «Tratar con dureza y con exigencias», «Provocar», «Hacer que un vehículo

[170] Cabe destacar que en su artículo la autora aborda dos tipos específicos de frecuencias: la absoluta y la relativa. Con frecuencia absoluta, la autora se refiere al número de apariciones de una locución en la red o en los corpus académicos «independientemente de si es mayor su aparición en textos orales que escritos» (Penadés Martínez, 2015: 255); la frecuencia relativa, en cambio, se relaciona con un aspecto específico de la lengua, por ejemplo, en función del tipo de texto en el que la locución es más habitual.

aumente de velocidad» y «Actuar con fuerza o con agresividad» (DILEA), por lo que el investigador o investigadora se ve obligado a realizar una revisión individual de cada locución.

Otro factor que la autora examina para describir la dificultad intrínseca en el cálculo de la frecuencia de las locuciones en un corpus es lo que define como «colisión homonímica», es decir, la homonimia que se produce entre una locución y una combinación de palabras no fraseológica. En este sentido, el sintagma *abrir la puerta*, puede ser fraseológico con el significado de hacer posible una cosa o hacer que ocurra (*abrir la puerta* + a algo/alguien) o, por otra parte, puede ser una simple combinación libre de palabras a partir del verbo *abrir*. Una vez más, el análisis crítico obliga a examinar locución por locución para evitar caer en el error. A partir de esta nueva perspectiva, la búsqueda de la frecuencia a través de Google (Sevilla Muñoz, 2012) se hace especialmente ambigua y metodológicamente complicada.

La tercera y última limitación que apunta Penadés Martínez es que, a menudo, la baja frecuencia en la utilización de ciertas locuciones va de la mano con la existencia de combinaciones de palabras no fraseológicas, homónimas de las locuciones (Penadés Martínez, 2015: 274). En otras palabras, una locución suele ser menos frecuente si su significante es el mismo que el de una combinación no fraseológica. Aunque no se trata necesariamente de una relación causa-efecto, es innegable de que existe una cierta relación entre la baja frecuencia y la homonimia de una locución con una combinación libre de palabras cuyo significado es diferente.

Como si todo eso no fuera suficiente, la autora menciona que, a diferencia de lo que pasa en otros idiomas como el inglés, donde se cuenta con datos sobre la frecuencia de las locuciones en grandes corpus como el BYC-BNC[171], «en el caso del español, los estudios que

[171] Véase Corpas Pastor (2013).

arrojen luz sobre esta cuestión en corpus de extensión similar son prácticamente inexistentes» (Penadés Martínez, 2015: 259). Dicha ausencia no hace más que complicar la tarea investigadora y, teniendo en cuenta la nefasta ausencia de materiales para la lengua italiana, es aún más ardua en ILE.

Las observaciones planteadas llevan a la investigadora a preguntarse qué criterios seguir a la hora de plantear un análisis de explotación de un corpus en función del criterio de la frecuencia:

> Es fácil imaginar la complejidad de dar respuesta exacta a estos interrogantes y la amplia casuística que se puede derivar de su aplicación potencial al conjunto de locuciones de una lengua. La tarea de su resolución parece titánica y el esfuerzo que conlleva puede ser descorazonador. De ahí la necesidad de tomar unas decisiones que, aunque provisionales y discutibles, faciliten la labor de redacción de un diccionario, la cual nunca se llevaría a cabo si hubiera que esperar a tener todas las locuciones completamente analizadas (Penadés Martínez, 2015: 271).

En este sentido, el investigador tiene que tomar posición frente a ciertas preguntas claves para poder identificar cuáles son las unidades fraseológicas más frecuentes para conseguir dicha finalidad. Será necesario, por tanto, escoger unos criterios previamente fijados como, por ejemplo, determinar el umbral mínimo de frecuencia de una unidad[172], comparar las UF escogidas con un corpus de referencia, o definir el grado de frecuencia de una unidad[173], etc.

[172] Entre otros criterios, la autora propone que: «la locución para la que no se hayan obtenido tres ejemplos podría o bien ser excluida de un diccionario de locuciones del español actual o bien ser marcada como poco frecuente [...] Por el contrario, tener ampliamente documentada una locución en relación con otras, después del vaciado de un conjunto de textos orales y escritos tomados como fuentes lingüísticas, puede ser criterio para marcarla como muy frecuente» (Penadés Martínez, 2015, 271-272).

[173] La autora propone emplear marcas como *rara*, y *poco frecuente*.

Teniendo en cuenta las anteriores propuestas, en el presente trabajo doctoral se considera que la didáctica de la fraseología de una lengua extranjera tiene que dar prioridad a las UF que muestran, en función de los criterios escogidos, una frecuencia más elevada. Además, si se pretende aplicar el corpus a un método didáctico, el criterio de la frecuencia tiene que plantearse en conjunto con otros criterios. Es, por lo tanto, necesario elaborar un listado de UF con fines fraseodidácticos que cruce distintos IF analizados en el presente apartado.

4.2.9. Los IF en breve

Tras analizar los principales indicadores propuestos por la academia española e italiana que sirven de puente entre la fraseología y su didáctica, es decir, entre la teoría y la práctica, se ha podido observar que:

a) El criterio de agrupación por léxico clave, a pesar de que haya sido uno de las más empleados para la elaboración de ejercicios didácticos, posee la importante limitación de priorizar el significante unitario de los componentes de una UF, frente a su significado composicional. Romper con esta premisa significa renunciar a una visión holística de la unidad, para reducirla al mero estado de léxico simple.

Asimismo, el riesgo de presentar distintas unidades bajo el enfoque unificador del elemento léxico reside en la fuerte relación sintáctica que se produce en la mente del alumnado. A pesar de que semánticamente las dos unidades sean diferentes, los estudiantes pueden confundir el significante de una con el significado de otra, o viceversa, creando unidades inexistentes. Los estudios analizados (Alessandro, 2011; Szyndler, 2015, etc.) demuestran que, pese a las similitudes que la fraseología presenta con el léxico, no es productivo abordar la fraseodidáctica bajo la misma perspectiva que la didáctica del léxico, en cuanto el significado unitario de una UF prevalece frente a sus componentes individuales.

b) Se ha observado que las propuestas para estudiar la fraseología muestran, con frecuencia, una tendencia a estructurarse en función de los campos semánticos. Dicha metodología puede favorecer la memorización de las UF si se analiza un conjunto en relación semántica de sinonimia, hiperonimia-hiponimia y antonimia y, por ende, las UF que derivan de esta metodología deberían de estudiarse/aprenderse como un conjunto único.

Asimismo, el concepto de sinonimia esconde ciertas dificultades que deben ser tenidas en consideración a la hora de elaborar una herramienta fraseodidáctica. En particular, merecen especial atención el aspecto pragmático, el registro al que pertenecen y la situación comunicativa.

c) El criterio del uso contrastivo es considerado, por lo general, un indicador válido a la hora de elaborar materiales fraseodidácticos por varias razones. En primer lugar, las lenguas presentan universales fraseológicos comunes en dos o más idiomas. En este caso, la descifración, memorización y el empleo de UF que de alguna manera se corresponden en la LE y la L1 será más fácil. En este sentido, cabe considerar como significativo dicho indicador, que se podría aplicar gradualmente al MCER: cuanto más parecidas sean las unidades entre la L1 y la LE, más básico será el nivel en el que se enseñarán. Así, la traducción se podría emplear también como recurso didáctico.

Llegados a este punto cabe señalar que dicha aplicación presenta ciertas limitaciones. En primer lugar, al trabajar con un par de lenguas exclusivas, se trata de un método de difícil replicabilidad en un aula lingüísticamente heterogénea, es decir, cuyo alumnado presenta diferentes L1. Si bien es cierto que en los países donde el italiano no es lengua materna las clases suelen estar compuestas principalmente por hablantes de la misma L1, esto no sucede en Italia, donde el alumnado habla distintos idiomas y, a menudo, muy lejanos sintácticamente del

idioma oficial del país donde se imparten las clases. A pesar de que dicha dificultad se podría subsanar aplicando el mismo método en todas las L1 presentes del aula, esta metodología complica notablemente su consecución práctica. Otro problema que presenta dicho indicador reside en los falsos amigos fraseológicos. Como se ha observado en profundidad en el presente capítulo, existen unidades que, en un par de idiomas dados, muestran similitudes en la forma, aunque el significado sea distinto. Frente a dicha situación, la presencia del profesorado o de un diccionario se vuelve imprescindible.

d) Un indicador que puede resultar interesante para el presente estudio es el de la agrupación por funciones comunicativas. Se trata de una herramienta valiosa dado que, por un lado, trata la lengua desde una perspectiva práctica e incidiendo en las relaciones pragmáticas que se instauran entre dos o más hablantes y, por otra parte, entre el contexto y el mismo sistema de la lengua. Desde este punto de vista, el corpus elaborado por Ana María Ruiz Martínez (2005) es de particular interés para la presente investigación.

No obstante, una vez, más el profesorado que decida emplear este indicador debe tener en cuenta ciertas limitaciones. Dichas dificultades tienen su origen en la traducción y la equivalencia dado que, a menudo, resulta complicado introducir un equivalente fraseológico de una LE por el simple hecho de que puede que no exista una determinada función comunicativa en la L1 de los aprendices. En este caso, será prioridad del profesorado insistir en esta diferencia pragmática y describir con otras técnicas las funciones discursivas que cada unidad desarrolla.

e) El criterio de la idiomaticidad está estrechamente ligado al de equivalencia. Cuanto más idiomática sea una UF, más complicado será para el alumnado descifrar la misma y encontrar un equivalente en su idioma materno. Dicha dificultad se debe a que la idiomaticidad de una unidad se ve influenciada por la manera de percibir la realidad por parte

de una comunidad de hablantes, la cual puede ser muy diferente a la de otra comunidad lingüística.

Por otra parte, el criterio de la idiomaticidad no puede tomarse en consideración, por lo menos como indicador exclusivo, a la hora de establecer una metodología fraseodidáctica que respete los niveles del MCER. Como se ha observado, unidades muy idiomáticas como *In bocca al lupo*, que resultan muy opacas para el alumnado de ILE, suelen ser enseñadas a partir de los niveles más básicos por su alta frecuencia en el uso. La idiomaticidad parece, por lo tanto, estar subordinada a otros indicadores como, por ejemplo, la frecuencia.

f) La motivación como indicador para elaborar un corpus de fraseologismos aplicados a la didáctica ha sido identificado por la academia como un criterio irrelevante. Si bien es cierto que la motivación de una UF puede suscitar el interés por parte del alumnado y despertar la emotividad, es subordinada a otros criterios primarios. La relación entre una unidad y sus orígenes etimológicos debería ser útil para analizar una lengua desde una perspectiva divulgativa, aportando información adicional sobre la unidad. Sin embargo, desde un punto de vista didáctico, es un mero dato añadido que se puede ofrecer *a posteriori* o de manera complementaria con el fin de suscitar curiosidad en el alumnado.

g) Finalmente, se puede afirmar que tanto el MCER como la gran mayoría de los estudios en didáctica de la fraseología consideran la frecuencia como el criterio predilecto para elaborar una colección de UF aplicable a la didáctica de una LE. En particular, numerosas investigaciones insisten en que, frente a una hipotética escala de frecuencia de las UF, el alumnado deberá, en primer lugar, adquirir un dominio estable de las formas más frecuentes y, posteriormente, dedicar sus esfuerzos a aprender las menos frecuentes. Cuanto más frecuente sea el uso una combinación de palabras, más alta será la posibilidad de que se cristalice

sin que se altere ninguno de sus constituyentes ni su forma. Desafortunadamente, para la lengua italiana no se dispone de un corpus basado en la frecuencia de las UF desde una perspectiva amplia que incluya todas las esferas de la fraseología, lo que lleva inevitablemente a que el profesorado -o cualquier investigador que quiera elaborar un corpus basado en dicho indicador- tenga que tomar decisiones cayendo en un cierto grado de subjetividad.

Por otra parte, los estudios analizados en el presente estado del arte han puesto en evidencia que el criterio de la frecuencia puede esconder trampas para la aproximación académica. Las propias características de las unidades fraseológicas, así como los fenómenos de polisemia y homonimia con las combinaciones de palabras libres, dificultan ulteriormente su extracción. Se ha observado que el criterio de la frecuencia es un dato de interés para elaborar un corpus solo si es acompañado por el análisis humano. Algunas expresiones idiomáticas pueden ser poco frecuentes desde el punto de vista estadístico, pero, aún así, forman parte del lexicón activo de un hablante nativo, al emplearlas correctamente tanto sintáctica que pragmáticamente y, además, identificándolas como parte integrante de su lengua.

En conclusión y por todo lo analizado en los apartados anteriores, la frecuencia de las UF se posiciona como un indicador fundamental a la hora de elaborar un listado de UF con fines fraseodidácticos, no debiendo aparecer las unidades con baja frecuencia en los niveles básicos. En este sentido, el/la investigador/a se ve obligado/a a tomar decisiones que, aunque provisionales y discutibles, faciliten compilar una colección de UF para la enseñanza/aprendizaje de una LE.

4.3. Principales propuestas metodológicas para enseñar fraseología como LE

A continuación, se presentan las principales propuestas metodológicas para enseñar fraseología italiana y española en las clases de lengua extranjera. Como se ha apuntado a lo largo del presente estudio, elaborar meras listas de vocablos y presentarlas al alumnado para que las aprendan de manera espuria y mnemónica es contraproducente para el aprendizaje activo (Alessandro, 2011).

Por ello, se debe plantear un criterio didáctico a través del cual presentar la fraseología en el aula. Dada la pertinencia y necesidad de esta reflexión, se ha planteado incorporar este análisis crítico al presente trabajo doctoral. A pesar de que no sea el objetivo principal de este trabajo, se pretende ofrecer una breve aproximación de las principales metodologías propuestas, si bien su análisis completo y la elaboración de un cuaderno de ejercicios específico basado en los resultados de esta tesis doctoral serán objetos de futuras líneas de investigación.

Cabe destacar que el presente estudio comparte con Olimpo de Oliveira Silva dos puntos de partida para entender el estado de la cuestión en esta materia. En primer lugar, que «más que propuestas didácticas, tenemos, en realidad, sugerencias que pueden usarse de manera puntual, pero no constituyen, de hecho, un trabajo que pueda desarrollarse a largo plazo» (2004: 125) y, en segundo lugar, que la mayoría de dichas sugerencias se llevan a la práctica en propuestas relacionadas con las locuciones, siendo esta la tipología de UF que más atención ha recibido en la enseñanza de una lengua extranjera ya que, debido a la idiomaticidad intrínseca que caracteriza gran parte de ellas, se ha considerado tradicionalmente como un escollo natural para el alumnado.

Asimismo, las propuestas que se analizarán comparten las consideraciones propuestas por Peter Kühn (1987). El autor propone el

término *fraseodidáctica* y, unos años después (1992, 1994, 1996), sugiere tres pasos o fases para incrementar el lexicón mental del alumnado de lenguas extranjeras. Kühn propone una primera fase de identificación, en la que el alumnado tiene que reconocer las unidades fraseológicas a partir de un texto real. La segunda etapa prevé que el alumnado, con la guía del profesorado, pueda descifrar las UF por sí mismo. Y, por último, la tercera fase se centra en la replicabilidad de las UF en un contexto real empleando un registro adecuado. Dichas consideraciones de Kühn servirán de inspiración para numerosas propuestas y la tricotomía reconocer-descifrar-reproducir se convertirá en la clave para futuras propuestas fraseodidácticas.

El análisis que se presenta en los siguientes apartados hará particular hincapié en las propuestas cronológicamente más recientes y se centrará en el ILE y el ELE, aunque de este último se proporcionarán más ejemplos, siendo su fraseología un terreno más analizado que el italiano. Finalmente, con el objetivo de presentarlas de manera homogénea, se han agrupado las propuestas en cuatro corrientes o acercamientos estructurales: contrastivo, semántico, funcional y misceláneas.

En el apartado 4.3.1. Contrastividad se presentan aquellas propuestas que se estructuran principalmente a partir una perspectiva contrastiva entre la LE y el idioma materno del alumnado. La siguiente sección, 4.3.2. Léxico, reúne aquellos enfoques que ponen el léxico al centro de la propuesta y, en particular, en torno a un determinado vocablo o un campo semántico: crisis económica (Mendoza Puertas, 2011), somatismos (Saracho Arnáiz, 2015), etc.

En el tercer bloque, 4.3.3. Funciones comunicativas, se presentan las propuestas metodológicas organizadas alrededor de las funciones comunicativas que las unidades expresan. En este sentido los autores ofrecen propuestas metodológicas para presentar al alumnado unidades que expresan, por ejemplo, valoraciones (Ureña Tormo, 2017) o alrededor

de las denominadas unidades fraseológicas de estampo pragmático (Alessandro, 2011). Por último, el apartado 4.3.4. Misceláneas ofrece todas aquellas propuestas que presentan técnicas mixtas y que, por ello, no tuvieron espacio en las secciones anteriores.

4.3.1. Contrastividad

En 1999, Penadés Martínez pública *La enseñanza de las unidades fraseológicas*, monografía que, por diferentes razones, marca un hito fundamental en futuros estudios sobre la aplicación didáctica de la fraseología. En primer lugar, se trata de un breve manual pensado para el profesorado de ELE en el que la autora ofrece de manera sistemática una clasificación y definición de los conceptos básicos de la fraseología, lo que es fundamental puesto que el profesorado desconoce el marco teórico de dichos estudios (Zuccalà, 2018). Por otra parte, Penadés Martínez sugiere al cuerpo docente manuales de interés para trabajar la fraseología en las clases de ELE. A pesar de que en su estudio la autora no proporcione ejercicios específicos, propone una metodología basada en tres fases: presentación, verificación y memorización de las UF en general y, en particular, de las locuciones.

Su metodología se basa en algunos conceptos claves que se refieren a los tres puntos mencionados, a saber: propone contrastar las unidades españolas con las de la lengua del discente explorando lo que define como «fondo cultural común» (Penadés Martínez, 1999a: 37). Es decir, desde un punto de vista contrastivo, analiza los que a lo largo del presente trabajo se han denominado como universales fraseológicos. Por otra parte, aconseja la redacción de un texto en el que se extrapolen las UF y se las sustituya por perífrasis. Y, finalmente, sugiere señalar las distintas acepciones de las unidades homónimas (literal y metafórica) y elaborar ejercicios de rellenar huecos.

Para reforzar la memorización de las locuciones, Penadés Martínez invita a utilizar el dibujo o la escenificación figurada de lo que expresa la unidad

o la redacción de una breve producción escrita utilizando las UF seleccionadas son solo algunas de las técnicas que Penadés Martínez propone emplear. También utiliza como ejercicio la sustitución de una frase de un texto por una UF equivalente en significado (relaciones de sinonimia, hiperonimia-hiponimia y antonimia). Es sus conclusiones, la autora lanza un reto a los investigadores que desean profundizar en fraseodidáctica y los anima a seguir estudiando dichos aspectos para que los resultados de sus trabajos puedan orientar la didáctica de las UF en el futuro.

Otro autor que dedica sus esfuerzos a la propuesta de una metodología contrastiva dirigida a la enseñanza de la fraseología en las clases de ELE es Gómez Molina (2000). En su trabajo sobre las locuciones idiomáticas, el investigador se basa en dos corrientes, la estructural y la cognitiva, y lleva a cabo un estudio empírico con estudiantes de ELE de nivel avanzado y superior. La concepción de análisis que emplea Gómez Molina incluye los valores pragmáticos y contextuales y, como se verá en los siguientes epígrafes, toma en consideración también los universales fraseológicos.

Su propuesta didáctica, pensada para todos los niveles de aprendizajes, pretende mejorar el proceso de enseñanza-aprendizaje de las UF otorgándole una importancia estratégica en el aula de ELE (Gómez Molina, 2000: 119). La base de su propuesta se inspira en el aprendizaje natural, es decir, que el alumnado aprenda fraseología como lo haría un hablante nativo: por un lado, desde un prisma contrastivo con su lexicón mental materno y, por otro lado, emprendiendo un recorrido autónomo a través de la fraseología.

Con el apoyo de los diccionarios fraseológicos, el alumnado podrá adquirir e implementar estrategias personales de aprendizaje que no son impuestas por el docente. Un dato interesante es que el autor propone orientar la elección de las unidades en función de «la frecuencia (intuitiva) de uso combinada con el criterio de eficacia (capacidad

semántica y dimensión pragmática)» (Gómez Molina, 2000:121). En este sentido, el diccionario se convierte en una herramienta imprescindible en su método.

A pesar de que este trabajo sea de gran utilidad práctica para el profesorado al demostrar que, como apunta el mismo autor, es posible motivar al alumnado y desarrollar su competencia idiomática de las locuciones, se considera que la hipótesis de base en la que se funda es errónea. Gómez Molina considera que las locuciones equivalentes o idénticas en la L1 y la LE son más fáciles de aprender, puesto que el estudiante dispone *a priori* de los códigos para descifrar la significación idiomática, la función expresiva y, por último, el uso sociolingüístico de estas locuciones. No obstante, el autor no considera que, como se apuntó anteriormente, muchos de los que aparentan ser universales fraseológicos son en realidad falsos amigos fraseológico y, por ello, aunque el estudio ha demostrado ser útil para el aprendizaje de los criterios sintácticos y pragmáticos, no propone una solución concreta con respecto a los criterios semánticos.

Con todo, cabe destacar que el método consigue forzar el alumnado a una reflexión para adentrarse autónomamente en los diccionarios. Como apunta Marello, la alfabetización digital puede empezar por los diccionarios y es:

> È fondamentale proporre esercitazioni che non siano tanto e solo per trovare delle parole di cui si ha bisogno nel momento, ma di far capire che l'informazione lessicografica organizzata, come è nei dizionari, è un modo per fare problem solving.[174] (Marello, 2018: entrevista en línea)

[174] Entrevista a Carla Marello para el Festival de las lenguas: Disponible en <https://www.youtube.com/watch?v=isWmhmHk21k>. «Es fundamental proponer ejercitaciones no tanto para encontrar las palabras que se necesitan en el momento, sino dejar comprender que la información lexicográfica organizada, como aparece en los diccionarios, es una manera para actuar la resolución de problemas» Traducción propia.

En 2006, Paula Quiroga pública el manual *Fraseología italo-española. Aspectos de Lingüística aplicada y contrastiva* en el que realiza un recorrido entre las principales clasificaciones fraseológicas en las dos lenguas objeto de estudio, elaborando un examen exhaustivo de la fraseología en los diccionarios monolingües y bilingües italiano-español, aplicando la fraseología a la traducción y, por último, dedicando un capítulo de conclusiones a la fraseología en la enseñanza del italiano como lengua extranjera. A pesar de que la autora no propone una verdadera metodología para impartir fraseología en el aula de LE, en el apartado 6.3. *Una nueva propuesta* (Quiroga, 2006: 198-200) ofrece algunas pautas que son de interés para esta investigación.

Quiroga pone de manifiesto que, a diferencia de lo que pasa con la lengua española, los estudios sobre fraseología italiana y su aplicación a la didáctica son muy escasos y, por ello, es necesario seguir investigando sobre dichos aspectos. Con respecto a una propuesta que tome en consideración el par de lenguas estudiado, la autora considera que en los niveles iniciales deben aprovechar la cercanía tipológica y cultural que se da entre los dos idiomas para enseñar/aprender fraseología; conforme el alumnado vaya adquiriendo un nivel más avanzado, se deberá que insistir en las diferencias. Por ello, considera que es necesario elaborar propuestas metodológicas que hagan hincapié en las similitudes y diferencias semánticas y sintácticas de las UF italianas y españolas puesto que, al ser dos idiomas con orígenes comunes, es fácil caer en errores de interferencia lingüística.

La autora propone elaborar dinámicas de ejercicios graduales en función de los niveles del MCER a través de tres fases: 1. identificación y comprensión; 2. reproducción y búsqueda de equivalentes; 3. desautomatización y comprensión de aspectos más pragmáticos que caracterizan la unidad (Quiroga, 2006: 199). En particular, la investigadora sugiere orientar los niveles más básicos hacia la identificación de la unidad (siguiendo los criterios de fijación y de significado no

composicional) y su comprensión. En el nivel intermedio, se debería hacer hincapié en la capacidad de reproducir las UF y en la búsqueda de unidades léxicas simples o perífrasis sinónimas y antónimas. Finalmente, en el nivel superior o avanzado se prestaría atención a la desautomatización de las unidades y a aspectos pragmáticos como el efecto humorístico de una UF o la posibilidad de una doble interpretación de la misma unidad, literal y fraseoló-gica.

En resumen, la propuesta de Quiroga se limita a ofrecer una pauta a seguir conforme el criterio de la equivalencia fraseológica en la enseñanza de las UF italianas a hispanohablantes en función de los niveles del MCER. La inserción de la fraseología deberá ser de manera gradual a partir del nivel inicial aprovechando la cercanía tipológica y cultural (equivalencia total) y en función del grado de conocimiento adquirido por el alumnado. Un aspecto interesante de su propuesta es que, a diferencia de Marta Higuera (1997), la autora no rechaza el aprendizaje de las unidades equivalentes ni de los universales fraseológicos.

En el mismo año, 2006, Rádis Baptista publica el estudio *Tratándose de expresiones idiomáticas, ¡no te rompas la cabeza ni busques cinco pies al gato!*, en el que analiza la manera en la que la información cultural de los fraseologismos está atendida en los diccionarios y en los manuales de aprendizaje. La investigadora de la Universidade Federal do Ceará aborda tres cuestiones principales. Después de ofrecer una definición de UF, analiza qué aporta al alumnado el aprendizaje de las expresiones idiomáticas y, finalmente, propone una metodología para facilitar su enseñanza y aprendizaje en el ELE para lusohablantes. Esta última parte será de particular interés para el presente estudio.

Con respecto al repertorio fraseológico que habría que seleccionar, Rádis Baptista adapta al contexto lusohablante el trabajo de Xatara (2001), en el que propone escoger las UF en función del grado de equivalencia con la LO. A partir de esa premisa, Baptista propone cuatro grados o niveles

distintos de dificultad de aprendizaje de UF. A saber: en el primer grado incluye las unidades que tienen equivalencia total con la lengua portuguesa puesto que suponen una menor dificultad en términos de enseñanza y de aprendizaje.

En el segundo grado, la autora brasileña selecciona las unidades que presentan cierta equivalencia en portugués, pero que no muestran cambios en la estructura, en el nivel de lenguaje, en el valor o en el efecto comunicativo. En el tercer nivel, selecciona las UF que se pueden traducir al portugués a través de otra unidad fraseológica distinta que no comparte la estructura sintáctica y/o unidades lexicales. Estas últimas presentan, según Baptista, mayor dificultad para su comprensión. Finalmente, el cuarto y último grado trata aquellas unidades que presentan equivalencia nula y, por ello, sugiere traducirlas por medio de paráfrasis.

Tras haber escogido la fraseología relavante, la autora se inspira en las pautas que Penadés Martínez ofrece en *La enseñanza de las unidades fraseológicas* (1999) para introducir el repertorio elaborado en la práctica didáctica. En particular, sugiere atender a cuatro etapas complementarias entre ellas: 1. presentación; 2. comprensión; 3. utilización y 4. memorización.

En la primera etapa de presentación prevé presentar las UF al alumnado en función del criterio de agrupación por palabras claves o por campos semánticos y sugiere que una manera para enseñarlas sería a través de preguntas sobre la descifración del significado en contexto solicitando ejemplos homólogos en la lengua de orígen. Con la segunda etapa, la comprensión, propone averiguar si el alumnado es capaz de reconocer y comprender las UF por sí mismo y sugiere emplear ejercicios de nexo entre unidades o de respuesta múltiple.

Con respecto a la tercera etapa, el uso, propone ofrecer ejercicios de rellenar huecos en los que el alumnado tenga que escoger entre las

actividades presentadas en las fases anteriores, o también ejercicios que lleven al estudiante a producir un texto escrito o oral en el que tendrá emplear una o más UF entre las propuestas. Finalmente, en la cuarta etapa, memorización, aconseja elaborar actividades de repaso del contenido estudiado y de las formas de las UF analizadas.

En conclusión, Rádis Baptista propone una metodología en dos macro etapas: en primer lugar, elaborar un corpus de UF siguiendo los criterios del grado de equivalencia para, en un segundo momento, presentar la selección fraseológica al alumnado en cuatro etapas consecutivas que llevan a la memorización de la UF.

4.3.2. Léxico

Alarcón Collado y López García (1992), aplicando las fases de Kühn previamente descritas (Reconocer-Descifrar-Reproducir), elaboran un listado fraseológico teniendo en cuenta elementos gramaticales, discursivos y sociolingüísticos, así como aportaciones de diversas metodologías de enseñanza. En este sentido, la fraseología es tratada como un aspecto discursivo homólogo al léxico y, como tal, recibe un enfoque práctico a la hora de desarrollar ejercicios.

En particular, las autoras consideran que se debe impartir fraseología, que denominan expresiones o modismos, de forma contextualizada a través de distintos fragmentos de la lengua hablada o escrita, y practicar dichas expresiones en cada uno de los seis niveles de competencia (del A1 al C2) a través de las cuatro destrezas: expresión oral, comprensión auditiva, expresión escrita y comprensión lectora. La principal aportación de las autoras en este sentido se da en el enfoque dirigido a la practicidad de la fraseología.

En 2011, Jorgue Daniel Mendoza Puertas pública en la revista RedELE (21) el artículo *De patitas en la calle y sin un duro: Una propuesta didáctica para la enseñanza de fraseología para un nivel B2*. En este artículo, el

autor pretende «integrar las unidades fraseológicas en el aula de ELE con un enfoque pedagógico más acertado» (Mendoza Puertas, 2011: 78). En particular, presenta una unidad didáctica desarrollada en el contexto de la crisis económica que estalló en España en 2008 y analiza la fraseología, locuciones sobre todo, que compone dicho campo semántico. En la primera fase de su propuesta, Mendoza Puertas sugiere empezar con la activación de conocimientos previos sobre el concepto de crisis y sus posibles declinaciones. A través de preguntas directas sobre el tema de la unidad se pretende preparar al alumnado para explorar y compartir de forma grupal sus ideas previas acerca del objeto del campo.

En una segunda fase, el análisis guiado de extractos periodísticos, llevará al alumnado a identificar distintas unidades fraseológicas, algunas más frecuentes que otras, así como a explorar su significado a través de una puesta en común en el aula. El profesorado tendrá que hacer hincapié en la idiomaticidad de las UF analizadas, con el fin de resaltar que algunas de ellas disponen tanto de un significado literal como composicional. Para fomentar la memorización de las unidades analizadas, se le propondrá a los estudiantes dos actividades: en la primera, el alumnado tendrá que unir la UF en contexto con su forma parafraseada; mientras que, en la segunda, habrá que distinguir cuáles de las UF utilizadas ofrecen un significado literal y cuáles figurado.

Con la segunda actividad de la unidad didáctica Mendoza Puertas establece asociaciones semánticas, formales e interculturales. Para lograr dicho objetivo, propone explorar expresiones sinónimas y antónimas, así como favorecer el vínculo entre «unidades de estructura y construcción sintáctica similar para agilizar y hacer más sencilla tanto la labor de asimilación de las mismas como las posteriores tareas de producción» (Mendoza Puertas, 2011: 83). Por otra parte, el alumnado también tendrá que practicar la aproximación contrastivo a través de la búsqueda de equivalentes en su idioma materno puesto que, según el autor, facilita el establecimiento de dichas redes.

A continuación, la actividad prosigue de la siguiente manera: se le propone al alumnado un nuevo texto para que busque, con el auxilio del diccionario, unidades fraseológicas sinónimas a las que se han analizado previamente en el otro extracto. Posteriormente, se procede a su agrupación estructural por similitudes (ej. tener narices - tener huevos - tener bemoles). El siguiente ejercicio tiene como objetivo, desde un prisma sintáctico, comprender si dichas unidades combinan con *algo*, con una cosa, o con *alguien*, con una persona, para después comenzar a trabajar sobre sinonimia a través de un ejercicio *ad hoc* en los que se propone al alumnado una serie de nuevas unidades en contexto, sinónimas o antónimas a las que se han analizado previamente, para que las conecte en función de su significado. Por último, se solicita al estudiantado que busque equivalentes en su lengua materna.

El último apartado de la unidad didáctica propuesta por Mendoza Puertas (2011) tiene que ver con el nivel pragmático de la lengua, el registro y el uso gramatical con el objetivo de ayudar al alumnado a evitar errores pragmáticos y emplear cada expresión en su correcta situación de interacción (formales, informales, muy informales, neutras y empleos tabúes)[175] En este sentido, el autor propone suministrar dos unidades, cuyo registro está claramente marcado y a partir de ahí, explicar los distintos registros de la lengua. En ese momento se solicitará al alumnado que, en grupos, relacionen las UF estudiadas con una de las situaciones comunicativas explicadas y, finalmente, se pondrán los resultados en común en el aula. La última fase de este apartado ofrece un ejercicio de rellenar huecos en una serie de diálogos para emplear las expresiones aprendidas que consideren más adecuadas. Finalmente, se le pedirá que elaboren situaciones comunicativas reales para utilizar

[175] Con respecto a los distintos registros de la lengua el autor sigue las marcas empleadas en el Proyecto de Excelencia de la Junta de Andalucía «Nuevo diccionario de aprendizaje (learner´s dictionary) del español como lengua extranjera de difusión online», dirigido por el Dr. Stefan Ruhstaller en la Universidad Pablo de Olavide de Sevilla.

algunas de las expresiones entre las estudiadas y que las escenifiquen en el aula a través de un *rol play*[176].

En 2015, Marta Saracho Arnáiz defiende su tesis doctoral[177] titulada *La fraseología del español: Una propuesta de didactización para la clase de ELE basada en los somatismos* en la que se aborda el panorama de las principales propuestas metodológicas para enseñar fraseología a través de un enfoque comunicativo, proponiendo una metodología fraseodidáctica. Su planteamiento está pensado para el alumnado del nivel B1, a través de nueve unidades fraseológicas: ocho locuciones y una fórmula rutinaria con dos variantes recopiladas en el *Diccionario Fraseológico Documentado del Español Actual* (Manuel Seco, Olimpia Andrés y Gabino Ramo, 2004). El criterio léxico es predominante en la elección de las nueve UF y, en particular, Saracho Arnáiz utiliza la agrupación por campos semántico, al ser unidades relacionadas con una parte del cuerpo humano.

Con respecto a la metodología empleada para elaborar las unidades didácticas, la autora presenta un método que define Modelo Metodológico del Heptágono (MMH) y que está basado en «una adaptación de la metodología de Gómez Molina (2004) para la enseñanza-aprendizaje del léxico, y en la metodología para la enseñanza aprendizaje del significado fraseológico de Timofeeva (2011)» (Saracho Arnáiz, 2015: 251). Dicha propuesta se divide en distintas fases cada una llevada a cabo a través de una o más actividades.

La primera fase se centra en los conocimientos previos del alumnado sobre los somatismos. Con la segunda fase, Marta Saracho Arnáiz propone presentar las unidades fraseológicas en un contexto real para, posteriormente, trabajar cuestiones de significado de los componentes

[176] En español juego de roles o interpretación de un papel. Dicha actividad didáctica supone dentro del aprendizaje que el alumnado se ponga a representar un papel asignado que refleje situaciones concretas de la vida real.

[177] Codirigida por Carmen Mellado Blanco y Tomás Jiménez Juliá.

léxicos que componen las UF y, sucesivamente, explotar el significado de la unidad en su totalidad. En otra fase sucesiva, el modelo plantea el análisis contrastivo con la lengua materna del alumnado; donde la producción activa de las UF está pensada para la fase de fijación de las UF en el lexicón del alumnado. Finalmente, la actividad de autoevaluación cierra la unidad didáctica.

Si bien la mayoría de las actividades propuestas por Saracho Arnáiz están pensadas para un trabajo individual, cabe destacar que existen otras elaboradas para ser realizadas en grupos o en pareja. La unidad didáctica ideada para el desarrollo de la competencia fraseológica llega a un total de 5h 45 min, por lo que se entiende que, desde desde un punto de vista meramente práctico considerando la dinámica del aula, las actividades que la autora detalla pormenorizadamente se tienen que desarrollar en varias sesiones.

Un aspecto de particular interés de esta propuesta didáctica es que ha encontrado aplicación directa en las nuevas tecnologías, a través de una página web educativa para la enseñanza y aprendizaje de ciertos somatismos: *Vive las expresiones sin pelos en la lengua*[178]. En ella, se utilizan diferentes herramientas y aplicaciones digitales como formularios de Google, Wikispaces, HotPotatoes, etc. Lo más interesantes de las actividades *on line*, además de su gratuidad, es que se trata de herramientas interesantes también para el alumnado autodidacta, que encontrará en ellas un valioso apoyo para desarrollar las destrezas lingüísticas activas, como la expresión escrita y la oral, frente a las pasivas de comprensión.

Por todo ello, la propuesta de Saracho Arnáiz adquiere relevancia en el panorama fraseodidáctico al encontrar dos grandes aplicaciones: por un lado, propone materiales y actividades pensadas para su uso en el aula

[178] Disponible en la dirección <http://martasaracho79.wix.com/vivelasexpresiones>.

de manera presencial y, por otro lado, ofrece recursos pensados exclusivamente para la red, a través del uso de aplicaciones digitales.

4.3.3. Funciones comunicativas

Arianna Alessandro[179] (2015) pública *Didáctica de la fraseología: una experiencia desde el marco de la investigación en la acción educativa*. Dicho ensayo fraseodidáctico, que parte de su investigación doctoral (Alessandro, 2011), dirigida por Pedro Guerrero Ruiz y por Pablo Zamora Muñoz, se centra en las unidades fraseológicas pragmáticas (UFP), subclase específica de fraseologismos, ya identificada por Zamora Muñoz (2003; 2005a). La investigadora presenta una propuesta didáctica[180] en la que, a partir de un modelo definido como Modelo Fraseológico Pragmático (Mo.Fra.P.), «plantea abordar la enseñanza y el aprendizaje de las UFP en el aula de E/LE e I/LS dentro del marco de la investigación-acción» (Alessandro, 2015: 173).

Esta obra es uno de los escasos trabajos en los que se abordan las unidades fraseológicas pragmáticas (UFP), a partir de un enfoque sistemático que combina la investigación y la didáctica. Las UFP, unidades con carácter de enunciado completo, se emplean en relación con el contexto social y la situación comunicativa y, por lo general, se utilizan para expresar sentimientos y estados de ánimo. En su mayoría, presentan una idiomaticidad reducida: son enunciados transparentes siendo semiidiomáticos o incluso literales. Por otra parte, poseen un grado de estabilidad menor que las locuciones porque están más sujetas a modificaciones o alteraciones por parte de los hablantes.

[179] La autora es miembro de FRASEMIA, grupo de investigación perteneciente a dicha institución y coordinado por Pablo Zamora Muñoz.

[180] Dicha propuesta se basa en una experiencia empírica llevada a cabo en el bienio académico 2007/09 teniendo como sujetos y destinatarios dos grupos de estudiantes, uno formado por italohablantes (ocho participantes) y otro por hispanohablantes (diez participantes) de tercer curso de titulaciones pertenecientes a la rama de Humanidades, respectivamente de la Università degli Studi Roma Tre (Italia) y de la Universidad de Murcia (España) (Alessandro, 2015: 180).

Su Modelo Fraseológico Pragmático (Mo.Fra.P.) aborda un enfoque «que privilegia la dimensión pragmática y el uso concreto de la lengua en situaciones de interacción verbal cotidianas e informales, con especial atención a aquellas que requieren el empleo de un registro coloquial» (Alessandro, 2015: 183). Se trata de una propuesta didáctica que se asienta en aspectos que caracterizan la forma y el uso de las UFP y, por medio de su explotación, busca el tratamiento didáctico más oportuno para el aprendizaje de estas unidades. Dicho modelo trata de abordar las UFP desde distintos tipos de contenidos, a saber: teóricos, fonéticos, gramaticales, léxicos-semánticos, funcionales, socio-culturales y desde una perspectiva contrastiva traductológica. En este sentido, el modelo propuesto por Alessandro sirve como base metodológica desde la cual abordar en el aula la enseñanza de las unidades objeto de estudio.

Su metodología se desarrolla en cuatro fases (Alessandro, 2015: 181 y ss). La primera fase denominada «diagnóstico y reconocimiento de la situación inicial», sirve al profesorado para comprender el nivel general del aula con respecto a la fraseología. En la segunda fase, «desarrollo de un plan de acción críticamente informado para mejorar aquello que está ocurriendo», la autora propone emplear el modelo del Mo.Fra.P. con dos finalidades: facilitar el aprendizaje del alumnado (finalidad empírica) y para que sea una herramienta en fraseodidáctica para futuras líneas de investigación (finalidad teórica).

Con la tercera fase, «actuación y observación», Alessandro propone llevar a la práctica, en el aula, el Mo.Fra.P, elaborando una serie de unidades didácticas basadas en una modificación de la secuencia didáctica propuesta por Woodward (2001): presentación de contenidos, práctica controlada y producción libre. Como resultado, presenta una serie de ejercicios divididos en cuatro secuencias, a saber: 1. Exposición a la lengua; 2. Percepción de la forma y el significado; 3. Memorización y uso; 4. Recapitulación y consolidación.

La autora propone acompañar a la parte práctica en el aula, caracterizada por la implementación de las secuencias didácticas elaboradas, y la recolección de datos para su posterior análisis a través de tres cuestionarios: uno previo a la actividad en el aula, uno intermedio, y uno al final de la puesta en práctica. En la cuarta y última fase de «reflexión y replanificación», Alessandro propone recolectar todos los datos obtenidos en las etapas anteriores para analizarlos y elaborar nuevas acciones en las que se mejoren los errores y las deficiencias del proceso didáctico, con el fin de ir perfeccionando las actividades.

En conclusión, se puede afirmar que el trabajo de Alessandro es innovador al tratar un aspecto de la fraseodidáctica poco estudiado en la lengua española y aún menos en la italiana, como son las unidades fraseológicas pragmáticas. La metodología muestra cómo compaginar la investigación en didáctica de la fraseología con su puesta en práctica en el aula para que cada docente pueda realizar este trabajo de manera individual convirtiéndose en agente formador y formado a la vez.

Por otra parte, cabe destacar que la autora emplea un corpus elaborado *ad hoc* en el que la selección y la introducción en la programación didáctica se limita a su mera experiencia subjetiva. Dicha forma de actuar, como ya se ha señalado, lleva a elaborar un corpus efímero y no holístico, diseñado con un alto grado de subjetividad y basado exclusivamente en la experiencia personal. En palabras de la autora:

> la selección y compilación del corpus está a cargo del docente el cual, desde su conocimiento de la lengua extranjera objeto de estudio, debe seleccionar las secuencias adecuadas para ser introducidas en su programación didáctica. (Alessandro, 2015: 184)

En 2017, se publica otra aportación interesante en la que se ofrece una secuencia didáctica desde el prisma de las funciones comunicativas para el aprendizaje de la fraseología en general y las locuciones en particular, en las lenguas extranjeras. A partir de un artículo de Ureña

Tormo (2014), Gómez González y la misma Ureña Tormo elaboran en una adaptación al ELE de dicho ensayo en la que proponen una metodología dirigida a estudiantes de español de nivel B2. El estudio examina un total de 17 locuciones de nivel B2, extraídas del estudio de Penadés Martínez (2015), «cuyo objetivo principal es que el discente –en este orden– identifique, comprenda, memorice y sepa utilizar el grupo de locuciones propuestas» (Gómez González y Ureña Tormo, 2017: 315).

En particular, las autoras plantean una secuencia que llaman «Valoramos como locos» en la que proponen actividades didácticas para trabajar sobre las locuciones en particular, que expresa la función comunicativa de la valoración en español. En otras palabras, las autoras se basan en el criterio de la agrupación de las UF por funciones comunicativas analizando unidades que expresan una valoración como:

- Me parece / Veo / Encuentro / Considero / Me resulta + adjetivo / SN... o + que + subjuntivo *(ej.: Me parecería muy bien que aprovecharas esa oportunidad)*.

- Lo veo / Lo encuentro + adjetivo / SN *(ej.: Todavía no quieren hacer pública la noticia, yo lo encuentro normal, ¿tú no?)*. (Gómez González y Ureña Tormo, 2017: 316-317)

A través de un enfoque explícitamente comunicativo, las investigadoras proponen una actividad de dos horas divididas en cuatro fases[181]. En la primera etapa, denominada «de calentamiento y de activación de contenidos» (Gómez González y Ureña Tormo, 2017: 317), se realiza un acercamiento al tema para motivar el alumnado y permitirle conectar sus conocimientos previos con los que se adquirirán. En esta fase inicial, el profesorado presenta una introducción a la fraseología para despertar la que Solano Rodríguez (2007) define como conciencia fraseológica, es decir, «que el aprendiz tome conciencia de la existencia y de la

[181] Cada fase prevé un tiempo diferente.

particularidad de las unidades fraseológicas, en tanto en cuanto universal lingüístico» (Rodríguez, 2007: 204-206). La segunda etapa prevé una comprensión lectora para reconocer los contenidos previamente dados.

A continuación, se da paso a la fase de la práctica controlada, en la que el alumnado ya puede empezar a emplear las locuciones que ha ido aprendiendo a lo largo de la tarea asignada. A través de un texto extraído a partir de un contexto de habla real, se presentan al alumnado nuevas locuciones en relación de sinonimia y de significado con las anteriores, para que este tenga que identificarlas y relacionarlas entre sí. Dicha etapa termina con la recopilación de las locuciones observadas y un análisis crítico en el que, por un lado, se identifican qué unidades expresan una valoración positiva y cuáles transmiten una negativa y, por otro lado, identifican si son empleadas para identificar de personas, de lugares, de objetos, de hechos y experiencias o si se usan de manera polivalente.

Finalmente, en la última fase denominada fase de práctica semilibre o libre, el alumnado tendrá que poner en práctica cuanto aprendido en el aula de forma oral y, trabajando en pareja, deberán contar uno de sus últimos viajes o hechos vividos. Las autoras, para finalizar la secuencia, proponen promover un enfoque contrastivo de las UF españolas e inglesas con el objetivo de señalar sus semejanzas y diferencias. En este sentido, la presencia del profesorado y la utilización de un diccionario bilingüe serán fundamentales para el éxito de la actividad.

4.3.4. Misceláneas

En un artículo de inspiración *kühniana*, Sevilla Muñoz y González Rodríguez (1995) proponen una serie de ejercicios pensados para realizarse de manera progresiva. Con el objetivo de que la enseñanza de una unidad fraseológica pueda tomar forma práctica en la didáctica, proponen

cuatro pasos: 1. Explicar el origen y significado de la UF (la motivación); 2. Presentarla en un contexto práctico de uso; 3. Contrastarla con expresiones sinónimas/antónimas y, por último, 4. Indicar su registro y frecuencia, en el caso de que se disponga de datos significativos.

Asimismo, proponen estructurar las actividades que ofrecer en el aula en tres bloques con metodologías variadas: ejercicios dirigidos a la descifración de la unidad; actividades para practicarla en contexto y dinámicas pensadas para su memorización. Será el alumnado, a través de fichas personalizadas, el que contribuirá a la creación de su propia base de datos de locuciones, anotando en estos materiales específicos información relevante como el significado de la unidad en su idioma, un ejemplo de uso, el registro al que pertenece, etc.

Otro trabajo interesante es el de Nuñez-Román (2015) quien propone «analizar la presencia de la fraseología en los niveles de enseñanza obligatoria y proponer una serie de actividades cuyo objetivo sea desarrollar la competencia fraseológica de hablantes nativos de español» (Nuñez-Román, 2015: 153). En este sentido, dicha actividad ha sido elaborada pensando en estudiantes de primaria, cuyo idioma nativo es el español, para que profundicen en su proprio idioma. A continuación, se examinará críticamente dicha propuesta ya que, a pesar de que esté pensada para hablantes nativos, se considera que algunas observaciones pueden ser relevantes para el alumnado de una LE.

Partiendo de la premisa ideológica de que «la enseñanza de la fraseología en lengua materna debe compartir las mismas estrategias y actividades propias de la enseñanza del léxico» (Nuñez-Román, 2015: 160), sin limitar su aprendizaje al mero almacenamiento mnemónico, Nuñez-Román elabora una propuesta metodológica tomando como punto de partida las observaciones de Josefina Prado Aragonés en su artículo *Didáctica de la lengua y la literatura para educar en el siglo XXI* (2004) y articula cinco estrategias para la adquisición del léxico, a saber:

1. Seleccionar adecuadamente el léxico según las necesidades comunicativas básicas de los alumnos, la frecuencia de uso y los centros de interés habituales;

2. Aprender el nuevo vocabulario a través de situaciones comunicativas reales y contextualizadas;

3. Aumentar la especialización del vocabulario en niveles superiores;

4. Poner énfasis en las relaciones semánticas (sinonimia, antonimia, homonimia) y cognitivas (relaciones conceptuales) que pueda establecer el nuevo vocabulario adquirido;

5. Hacer uso de estrategias y actividades motivadoras y lúdicas, incluyendo actividades de refuerzo que permitan consolidar el vocabulario adquirido y permitan incluir el uso del diccionario en clase (Nuñez-Román, 2015: 161).

Por otra parte, el autor ofrece una propuesta de secuencia de actividades dirigida a desarrollar la competencia fraseológica en cinco fases: reconocimiento, fijación, exploración, explotación y refuerzo. En la primera etapa de reconocimiento se ofrece al alumnado una serie de unidades extraídas de un texto oral o escrito y serán los mismos estudiantes quienes, con el apoyo del diccionario, busquen su significado y lo anoten. Para la siguiente fase, fijación, Núñez-Román propone distintos ejercicios orientados a la manipulación e interiorización de las UF como, por ejemplo, dibujos explicativos de las unidades, *rol play* o actividades lúdicas como sopas de letras o crucigramas.

A través de la tercera fase de exploración, el autor plantea evaluar las relaciones semánticas o temáticas con otras UF y, para ello, diseña varias actividades como la búsqueda de otras unidades en función del criterio de agrupación por palabras claves, por temas o por sinonimia. En la cuarta etapa, explotación, el estudiante tendrá que emplear las UF aprendidas en contextos de uso adecuados. Para ello, Núñez-Román

propone ofrecer ejercicios en las que se sustituyan palabras simples por las unidades analizadas previamente en frases o extractos de textos, crear frases a partir de una unidad estudiada, o, finalmente, elaborar frases en la que la misma UF sea empleada de manera idiomática y literal. Por último, la fase de refuerzo servirá para fortalecer los conocimientos adquiridos a través de actividades, a partir de la reconstrucción de unidades por la mezcla de diferentes fraseologismos estudiados o la búsqueda de errores en el uso de las UF aprendidas.

El autor ofrece algunas pinceladas sobre cómo enseñar fraseología a alumnado de secundaria y, en este sentido, propone «el análisis detallado de tipologías textuales más complejas [...] agrupadas por contextos situacionales, por sus funciones discursivas y por su estructura gramatical» (Núñez-Román, 2015: 162). En este nuevo nivel de aprendizaje, se propone analizar textos periodísticos en los que el profesorado tendrá que resaltar aspectos específicos, como el uso frecuente de conectores fraseológicos, las numerosas locuciones adverbiales y preposicionales y las colocaciones y locuciones nominales.

Otro trabajo que representa una interesante aportación a la enseñanza de la fraseología del ELE[182] es el de Ureña Tormo en 2017. El artículo, que toma como punto de partida las mismas bases teóricas del que se han descrito críticamente con anterioridad al tratar el texto escrito con Gómez González, propone una metodología que evalúa 74 locuciones pertenecientes al nivel B2[183] y 52 refranes del español para el nivel C1, dirigidos a la enseñanza de español como lengua extranjera (nivel B2 y C1), cuya lengua materna es el inglés. En particular, la autora propone 89 actividades para los refranes y 42 para las locuciones divididas en «cuatro secuencias didácticas, en cada una de las cuales se trabaja uno de los

[182] El artículo es un producto científico resultado del Trabajo de Fin de Máster (2014).

[183] Las locuciones, así como los niveles del MCER al que pertenecen, han sido extraídas de las obras lexicográficas de Penadés Martínez (2002, 2005, 2008) y los refranes principalmente de Penadés Martínez y Olimpio de Oliveira (2008).

cuatro contenidos habituales: funcional, gramatical, léxico y cultural, junto con diferentes aspectos de las locuciones» (Ureña Tormo, 2017: 100).

Centrándonos específicamente en las locuciones, la autora propone dividir las cuatro fases de la secuencia de la siguiente manera. La primera está pensada para trabajar la función comunicativa a través de locuciones cuyo contenido funcional es el de la valoración desde un punto de vista contrastivo, con los equivalentes en inglés; la segunda etapa aborda, desde el nivel gramatical, los aspectos formales y de combinatoria sintáctica de los verbos ser y estar; la tercera fase está dirigida al campo semántico del viaje, y trabaja el aspecto semántico y el registro de uso; y, finalmente, la cuarta se centra en el campo semántico de la tauromaquia y aborda el aspecto del significado y el de la motivación de las locuciones (Ureña Tormo, 2017: 100).

Con respecto a los refranes, la autora propone ocho secuencias que incluyen los contenidos funcionales de la expresión de la advertencia, de valoraciones y creencias. Por otra parte, los refranes que presentan unidades gramaticales son empleados para analizar los pronombres relativos, mientras que los que se construyen sin verbos abordan las oraciones nominales y condicionales. Los contenidos léxicos son trabajados a través de la variedad mejicana[184] en los campos semánticos de la comida y de las personas. Por último, los contenidos culturales están relacionados con el campo de la religión, leyendas y la historia de España.

La autora propone solo dos tipologías de actividades de una secuencia didáctica, por lo que es imposible analizar críticamente la propuesta metodológica en su conjunto. Sin embargo, Ureña Tormo asegura que:

[184] Cabe destacar que no se pudo encontrar el TFM completo para elaborar la presente tesis doctoral. No obstante, en el artículo analizado no se justifican las razones que han llevado a introducir una variedad diferente a la peninsular, ni tampoco las razones que han llevados la investigadora a elegir la variedad de Méjico frente a otras.

> Las secuencias didácticas propuestas reflejan el planteamiento que se desprende del enfoque comunicativo de la lengua - y que - todas las actividades persiguen uno o varios de los siguientes objetivos: reconocer, comprender, memorizar y saber utilizar estas dos clases de fraseologismos mediante una combinación de distintas dinámicas de trabajo que fomenten el trabajo individual, por parejas y en equipo. (Ureña Tormo, 2017: 100)

Otra reciente aportación proviene de la tesis doctoral de Rosario Lisciandro (2018), *Las unidades fraseológicas en la enseñanza de una lengua extranjera (italiano - español)*. En el quinto capítulo de su investigación, Lisciandro propone seis unidades didácticas dirigidas a estudiantes de italiano como lengua extranjera, pertenecientes a casi todos los niveles que concierne el MCER (del A1 al C1). Las actividades han sido pensadas de manera progresiva, en función de los distintos tipos de unidades propuestas aunque la autora no describe pormenorizadamente los indicadores que le han llevado a forjar dicha gradualidad. Los ejercicios descritos se dividen en cuatro fases: una primera de motivación inicial; una segunda, que prevé ejercicios lexicales; a continuación, una actividad para desarrollar la expresión oral y una última fase de expresión escrita para fortalecer lo aprendido en aula.

La propuesta didáctica para los niveles del MCER A1 y A2 de Lisciandro se presenta de manera conjunta, aunque dicha agrupación puede comportar cierta crítica puesto que tanto los contenidos gramaticales, como los sintácticos son distintos para cada nivel[185]. En la propuesta, se abordan unidades pertenecientes al campo semántico de la vida cotidiana y, en particular, se tratan las colocaciones compuestas por V+N como *fare la doccia, preparare il pranzo*, etc. La segunda propuesta didáctica dedicada al nivel B1 presenta unidades fraseológicas

[185] Véase a respecto el Capítulo 8 del Marco Común Europeo de Referencia para las Lenguas: *La Diversificación lingüística y el curricolo* (Consejo de Europa, 2001: 167).

pertenecientes al ámbito hostelero (Lisciandro, 2018: 228) y, en particular, trabaja las colocaciones V+N y las locuciones nominales.

Con respecto al nivel B2, la autora propone tres propuestas didácticas diferentes. La primera plantea el estudio de las locuciones compuestas por los verbos *essere* y *avere,* como *essere in vena di* (algo), *essere l'ultima ruota del carro* o avere *culo*, *avere cervello*, etc. La siguiente se centra en las «fórmulas para redactar una carta formal, entre las cuales las de inicio y cierre» (Lisciandro, 2018: 244) y colocaciones nominales pertenecientes al campo semántico del trabajo como *busta paga*, *contratto di lavoro*, etc. La tercera y última propuesta aborda locuciones y refranes italianos caracterizados por una fuerte idiomaticidad como *Rimanere a bocca aperta*, *Scambiare lucciole per lanterne*, etc.

Lisciandro explica que el criterio elegido para seleccionar las UF es el de la frecuencia: «las unidades fraseológicas que componen los ejercicios a causa de la frecuencia de uso en la lengua italiana, sobre todo en la lengua hablada» (Lisciandro, 2018: 252). Con todo, no da mayor información sobre los indicadores empíricos que le han llevado a definir la que identifica como «frecuencia de uso la lengua italiana», por lo que cabe suponer que dicha elección ha sido afectada por cierta subjetividad. La última propuesta didáctica que elabora está dirigida al nivel B2 y C1 del MCER, y trata ciertas «unidades fraseológicas con sentido metafórico» (Lisciandro, 2018: 258) y, en particular, las locuciones verbales con los verbos *andare, avere, dare, essere* y *fare* como *andare con i piedi di piombo*, *dare filo da torcere*, etc.

En conclusión, las propuestas didácticas elaboradas por Lisciandro (2018) presentan algunas limitaciones pero, al mismo tiempo, articulan un avance significativo en el panorama de le fraseodidáctica del ILE. Con respecto a las debilidades de su investigación, cabe subrayar que no se ofrece un verdadero método para facilitar al profesorado la posibilidad de elaborar sus propias unidades, sino se presentan actividades

específicas, principalmente para el nivel B2, relacionadas con una determinada unidad didáctica o un campo semántico específico.

Sin embargo, es indiscutible que el estudio es de particular interés por cuatro razones principales: en primer lugar, se trata de una investigación necesaria y relevante, al ser uno de los pocos trabajos enfocados hacia la fraseodidáctica del italiano como lengua extranjera dirigido a hispanohablantes que presenta un compendio práctico de actividades y ejercicios; en segundo lugar, ofrece un interesante corpus fraseológico extraído de 25 manuales de italiano como LE, relativamente recientes; además, las actividades propuestas han sido experimentadas en centros de enseñanza del italiano como lengua extranjera, tales como la Escuela Oficial de Idiomas de Almería, la Universidad de Almería, la Universidad de Córdoba, la Universidad de Innsbruck (Austria) o la Università per Stranieri di Siena (Italia), lo que significa que han pasado por un filtro antes de ser publicadas; y finalmente, el trabajo aborda todo el universo fraseológico, frente a otros que se centran exclusivamente en una esfera determinada.

El mismo año, Angela Mura publica un nuevo artículo titulado *Fraseología y fraseodidáctica digitales en E/LE* (2018) cuyo objetivo es explorar los recursos que las Tecnologías de la Información y la Comunicación (TIC) ofrecen para trabajar la fraseología dentro y fuera del aula, y plantea una propuesta didáctica en forma de *webquest*. En la primera parte del ensayo, la autora presenta una serie de herramientas y recursos *on line* necesarias para estudiar la fraseología y las divide en cinco categorías 1. diccionarios electrónicos como el DLE (Diccionario de la Lengua Española, 2004), el DEC (Diccionario del Español Coloquial, 2012), el DICE (Diccionario de Colocaciones del Español) (2004) y El Refranario (Diccionario de refranes en español, 2012); 2. corpus en línea entre los que cita el CREA (Corpus de Referencia del Español Actual), El COLA (Corpus Oral de Lenguaje Adolescente) y el C-ORAL-ROM (Corpus Oral de Lenguas Románicas); 3. páginas web especializadas como PAREFRAS

(Paremiología y Fraseología), FRASYTRAM (Fraseología y Traducción Multilingüe), PHRASEONET (El mundo de las frases hechas) y FRASEMIA (Fraseología, Paremiología y Traducción); 4. revistas digitalizadas, solo menciona la revista en línea Paremia creada en 1997 por el Grupo PAREFRAS; y 5. bases de datos como El Refranero Multilingüe (2005), BADARE (2008) y FRASYTRAM (2007).

En la segunda parte de su artículo, la autora profundiza en cómo «estas herramientas pueden ayudar a un profesor de lenguas en el proceso de enseñanza de las UFS» (Mura, 2018: 252). En particular, presenta una actividad basada en una herramienta *on line* específica: la *webquest,* es decir, una actividad guiada en la que se propone la resolución de una tarea al alumnado y este, de manera colaborativa y cooperativa, tiene que buscar la información necesaria en la red. La actividad, denominada FraseoEle (Fraseología, Español Lengua Extranjera), está pensada para un nivel C1 y se ha realizado a través de Blogger, el servicio de Google.

A pesar de que el acceso a dicha herramienta sea restringido[186], la autora asegura que las tareas elaboradas están pensadas con el objetivo de promover las competencias léxicas, semánticas y ortoépicas de las UF. La *webquest* se articula en ocho apartados: después de una introducción en el que el profesorado, a través de un vídeo, acerca el alumnado al objeto de estudio, se le plantea la tarea final: crear un blog de fraseología española en el que recopilar y registrar los logros conseguidos. A continuación, se plantean las siete tareas que los estudiantes tendrán que realizar para alcanzar el ya descrito ejercicio final. Dichas actividades «se realizarán en grupo de acuerdo con la siguiente dinámica: la clase se divide en tres grupos; todos los grupos harán las primeras dos actividades y las dos últimas, mientras que las tres actividades intermedias se distribuirán en los tres grupos» (Mura, 2018: 255).

[186] Desafortunadamente, no se puede acceder a la *webquest* sin invitación. En febrero de 2019, el investigador lo solicitó a través de la página pero nunca se no tuvo respuesta, por lo que no se ha podido profundizar en el análisis. La dirección de la webquest es la siguiente <https://fraseoele.blogspot.com/>.

La primera actividad pretende acercar el alumnado al objeto de estudio. Por ello, se le ofrecerá la clasificación que Corpas Pastor (1996) elabora para describir el universo fraseológico y se le pedirá que encuentren ejemplos de cada una de las Esferas descritas por la autora en la trascripción de un monólogo de Enrique San Francisco (*Los Velatorios*[187]). La segunda actividad prevé la búsqueda de las unidades encontradas en los diccionarios en línea correspondientes: las locuciones en el *Diccionario del español coloquial*, el *Refranario* para los enunciados fraseológicos y el *Diccionario de colocaciones del español* para la esfera de las colocaciones.

Las siguientes actividades (3, 4 y 5) están pensadas para tres grupos distintos, con el objetivo de que cada uno trabaje sobre de manera individual en una de las tres Esferas: el grupo «locuciones» hará la actividad 3 «Piensa», el grupo «enunciados fraseológicos» se dedica a la actividad 4 «Juega», mientras que el grupo «colocaciones» a la actividad 5 «Colócate» (Mura, 2018: 256-257). La sexta actividad se desarrolla fuera del aula y en colaboración con alumnado nativo de lengua española. Los estudiantes de ELE tendrán que descubrir y grabar con el móvil los gestos que se asocian con unas unidades específicas españolas. La última actividad, la séptima, plantea la tarea final: la creación del blog de fraseología española en el que los participantes introducirán toda la información recopilada.

A pesar de que, como se ha apuntado, no se pudo acceder a la web y experimentar en primera persona las actividades, la metodología propuesta resulta interesante para el presente trabajo por distintos motivos. En primer lugar, se considera que los medios digitales son los predilectos para el alumnado, en cuanto despiertan su atención e interés al recibirlo como una novedad frente al libro de texto. Por ello, el profesorado tiene que esforzarse en utilizar las TIC en sus clases, tarea

[187] Diponible en <https://www.youtube.com/watch?v=-_BVISWZ_OU>.

cada día menos ardua, gracias a la gran cantidad de contenido ludo-didáctico en la red. Por otra parte, la actividad que trabaja el aspecto kinésico de la fraseología (actividad seis), lo que resulta particularmente interesante si se aplica a la lengua italiana, donde se hace gran uso de la narración gestual y del aspecto extralingüístico como elemento enfático o donde incluso asume propiedades lingüísticas en situaciones donde el habla no es posible.

Por último, cabe destacar en este apartado las aportaciones del reciente proyecto GEPHRI a la didáctica de la fraseología italiana. El proyecto GEPHRI (*GEbrauchsbasierte PHRaseologie des Italienischen*)[188] nace en 2018 por mano de Elmar Schafroth y Riccardo Imperiale con el objetivo de llevar los principios de la gramática de construcciones (CxG del inglés *Construction Grammar*) a la fraseología del italiano. El resultado es una plataforma en línea que presenta más de 300 locuciones verbales frecuentes y actuales de la lengua italiana. La aportación más valiosa de dicha herramienta en línea es que por cada UF se presenta de manera sistemática y pragmática (de manera similar a un diccionario) la información sobre cuatro niveles lingüísticos: semántico, sintáctico, pragmático y didáctico.

Con respecto al nivel semántico, se define el significado composicional y el campo semántico al que cada UF pertenece; en el nivel sintáctico, en cambio, se definen las posibles variantes de una locución (incluidas las menos comunes) y los colocativos de la unidad, es decir aquellos componentes léxicos que se pueden/suelen insertar al principio, al final o en el medio de la UF; por otra parte, el nivel pragmático se centra sobre el uso, el fin comunicativo y el registro de la unidad en cuestión, información determinante para evaluar y comprender el acto ilocutivo de una UF. Finalmente, en el nivel didáctico, especialmente enfocado hacia el hablante germanófono, se presenta rico de información como

[188] Para más información sobre la plataforma fraseológica véase < https://gephri.phil.hhu.de/>.

indicaciones de uso, precauciones con respecto a las connotaciones estilística, falsos amigos y videos que muestran el uso de la UF en la lengua oral. En conclusión, dicho proyecto demuestra sin duda un avance significativo en cuanto a la manera de abordar la fraseología desde una perspectiva fraseodidáctica en cuanto proporciona al alumnado valiosa sobre la comprensión y el uso de la fraseología italiana.

4.3.5. Propuestas fraseodidácticas en breve

Para concluir este capítulo, se explicarán los puntos de intersección que se dan entre las diferentes propuestas metodológicas analizadas críticamente, con el objetivo de profundizar en aquellos aspectos comunes que se presentan desde un prisma ecléctico.

a) En primer lugar, se observa que, más que un verdadero modelo holístico a seguir para orientar un trabajo a largo plazo, las diferentes propuestas didácticas son, sobre todo, sugerencias dirigidas a actividades específicas y puntuales. En este sentido, la gran mayoría no abarcan todo el espectro didáctico, sino que se centran en la explotación de algunos aspectos particulares: una unidad temática en particular, un campo semántico específico (Lisciandro, 2018), ciertas unidades fraseológicas relacionadas con unas palabras claves (Núñez-Román, 2015), etc. Se observa, por lo tanto, una ausencia de un modelo universal y replicable que funda la consolidada clasificación del universo fraseológico con la la didáctica de una lengua extranjera.

b) Dicha falta de una propuesta holística se ve corroborada por el hecho de que la mayoría de las unidades didácticas propuestas priorizan las locuciones, dejando las otras esferas del panorama fraseológico en un plano marginal. Ello permite corroborar que nos encontramos frente a la ausencia de un modelo fraseodidáctico global que involucre, desde un punto de vista amplio, todo el espectro que alcanza la fraseología. Este es el principal escollo que la fraseodidáctica tendrá que superar y el principal reto para futuros estudios

c) Por otra parte, la mayoría de las metodologías fraseodidácticas analizadas se basan las tres fases propuestas por Kühn (1992, 1994, 1996) que desarrollan el grado de dominio de la fraseología aprendida de manera progresiva: identificación de la UF, interpretación y comprensión de la misma y, finalmente, puesta en práctica de las unidades aprendidas en en un contexto real. Para ello, de acuerdo con Marta Báralo (2006), se requiere que el alumnado desarrolle ciertas capacidades cognitivas de retención, fijación y reutilización mediante ejercicios que requieran esfuerzos de tipo comunicativo. En este sentido, las unidades didácticas analizadas son concebidas de forma progresiva, desde el nivel A1 hasta el nivel B2/C1 y suelen repetir el mismo patrón: una primera fase de motivación inicial, ejercicios para la descifración, actividades de expresión oral y de expresión escrita para su puesta en práctica desde un punto de vista comunicativo.

d) La destreza comunicativa predomina en todas las propuestas analizadas. La gran mayoría de las metodologías comparten las hipótesis de Corpas Pastor (1996, 2001, 2013) y de Penadés Martínez (2015), según las cuales el dominio de las unidades lingüísticas genera en el alumnado una sensación de integración en una comunidad de habla. A este fenómeno se añade el factor motivacional de que la fraseología estimula y despierta curiosidad en los estudiantes. Por todo ello, el profesorado debe aprovechar este impulso y promover el estudio de las unidades fraseológicas en el aula.

e) Los trabajos analizados críticamente ponen de manifiesto que la fraseología debe ser enseñada teniendo en cuenta su estrecha relación con el resto de los contenidos del currículo y, asimismo, reflejar los aspectos funcionales, gramaticales, léxicos y culturales de un idioma (Penadés Martínez, 2015; Ureña Tormo, 2017). En otras palabras, la enseñanza de la fraseología tiene que estar incluida en todas las unidades didácticas, sirviendo como complemento a los manuales que, en muchas ocasiones, demuestran trabajarlas de una manera dispersa y

aleatoria, independiente al resto de contenidos académicos (Olímpio de Oliveira Silva, 2006; López Vázquez, 2010; Gómez González y Ureña Tormo 2017). Para un desarrollo completo de las habilidades del alumnado, la enseñanza de la fraseología tiene que ser integrada con el resto de contenidos que se incluyen en los manuales, vinculando las UF a todos los contenidos que se van a tratar.

f) Por último, algunas propuestas coinciden en que adoptar un enfoque contrastivo entre la fraseología de la lengua extranjera y la correspondiente en lengua materna del estudiante puede ser realmente beneficioso para el alumnado. Según Leal Riol (2011), la adopción de este enfoque resulta indispensable para la total adquisición de las unidades fraseológicas. En este sentido, numerosas son las propuestas que ponen el acento en el carácter universal de la fraseología en general y de las locuciones en particular, puesto que resulta muy útil y provechosa para los procesos de la enseñanza-aprendizaje.

CAPÍTULO V

METODOLOGÍA EMPLEADA

5.1. Introducción

Con el objetivo de estudiar los que a lo largo de este trabajo se han definido con el término «Indicadores Fraseodidácticos» (IF) aplicables a la didáctica de la fraseología dentro de las clases de italiano como lengua extranjera (ILE), se ha empleado un diseño metodológico de tipo cualitativo y cuantitativo. En particular, se ha recurrido al modelo secuencial exploratorio (Creswell, 2003, 2007, 2009, 2014).

Dicho modelo metodológico propuesto por el profesor estadounidense de la New Jersey City University, John Cresswell, trata de conseguir la mayor información sobre el objeto de estudio, en este caso, la enseñanza de unidades fraseológicas en las clases de italiano como lengua extranjera, mediante distintas técnicas aplicadas de manera secuencial. La aplicación continuada de distintos métodos permite avanzar en el diseño metodológico, de manera que los resultados de la primera fase permiten continuar con la segunda etapa para sucesivamente proceder con la tercera y así sucesivamente hasta finalizar el diseño con la última fase de la metodología.

El diseño multimétodo planteado por el profesor estadounidense se ha adaptado a las necesidades de la presente investigación mediante cuatro fases diferenciadas (Véase Figura 5) en las que se utilizan métodos cuantitativos y cualitativos con el objetivo de analizar la compleja

realidad de las unidades fraseológicas aplicadas a la didáctica en el aula de italiano como LE. En este sentido, la metodología del presente trabajo parte de una lógica inductiva que permite diseñar una propuesta metodológica acumulativa basada en cuatro fases consecutivas:

1. Revisión documental;
2. Extracción y anotación fraseológica;
3. Propuesta de análisis de la frecuencia en el *Focus corpus*;
4. Propuesta de aplicación de los indicadores fraseodidácticos al *Reference corpus*[189].

Figura 6. Diseño metodológico secuencial exploratorio empleado.



Fuente: Elaboración propia.

La primera fase contempla la revisión documental de fuentes primarias y secundarias. Para ello, se analizaron críticamente las principales propuestas metodológicas empleadas para enseñar fraseología italiana y española en el aula con el objetivo de ofrecer un estado de la cuestión holístico y poner los cimientos para futuras líneas de investigación. Por otra parte, se elaboró un estudio pormenorizado de las principales propuestas para trabajar la fraseología en el aula de italiano y español como lengua extranjera. En primer lugar, se elaboró un análisis crítico de los indicadores fraseodidácticos más productivos para organizar el material fraseológico analizando exclusivamente aquellas propuestas que se basan en criterios empíricos y omitiendo los que presentan distintos grados de subjetividad, es decir, los que encuentran sus raíces exclusivamente en la experiencia personal de su autor/a. Cabe destacar que, por primera vez, se sistematizan en un único trabajo los principales indicadores

[189] Elección de los términos *Focus corpus* y *Refence corpus* de acuerdo con la terminología de Sketch Engine (véase apartados 5.3.1. y 5.3.2.)

fraseodidácticos para elaborar listado de UF aplicado a la didáctica del italiano como lengua extranjera.

Tras la revisión documental, la segunda fase de la metodología prevé la extracción manual de la fraseología a partir dos series completas (A1-C2) de manuales de reconocido prestigio en el ámbito de la didáctica del italiano como lengua extranjera: *Nuovo Espresso* y *Il Nuovo Affresco Italiano*. Se optó por realizar una extracción manual frente a una automática principalmente por aquellas limitaciones que todavía presentan las herramientas de extracción terminológica y que se definirán en los siguientes apartados. En este sentido, se realizó un pre-test empleando tres herramientas de extracción automática, a saber: Sketch Engine de Lexical Computing Limited, SynchoTerm de Terminotix y, por último, el reciente Term Extraction Module (TEM) de IATE, Interactive Terminology for Europe. Sin embargo, los resultados no fueron satisfactorios por diferentes razones que se detallarán a lo largo del siguiente apartado.

Tras identificar y extraer la fraseología relevante para el presente estudio, se elaboró un archivo fraseológico[190] *ad hoc* en formato digital en el que se anotaron y transfirieron las UF. La creación de dicha herramienta se basó en las normas de la *International Organization for Standardization* (ISO)[191] y, en particular en la ISO 704: 2000, *Terminology work. Principle and methods* y en la ISO 1087-2: 2000, T*erminology work - Vocabulary - Part 2: Computer applications*. Dichos estándares fundamentales para la labor terminológica han sido aplicados al área de la fraseología.

La tercera fase prevé la comparación de las unidades fraseológicas extraídas (*Reference* corpus) con un corpus de lengua general del

[190] Con el término archivo fraseológico se entiende un archivo en-linea (formato .xlsx) que presenta un listado de UF ordenadas a partir de criterios comunes (véase apartado 1.5.).

[191] La ISO es una organización internacional independiente y no gubernamental fundada en 1946 y con sede en Ginebra. La ISO cuenta con 164 organismos nacionales de normalización y su finalidad principal es la elaboración de estándares internacionales, es decir, normas basadas en el consenso internacional y pertinentes para el mercado. Para más información véase https://www.iso.org/standards.html.

italiano (*Focus* corpus) con el objetivo principal de analizar el indicador fraseodidáctico específico de la frecuencia. Para ello se propone el empleo de una herramienta informática *ad hoc* creada en colaboración con el *IATE Support & Development Team*. Dicha herramienta permitirá contrastar las unidades extraídas manualmente con un corpus general de la lengua italiana compuesto por más de 4,9 mil millones de palabras y devolver el grado de frecuencia absoluta en la lengua general.

Finalmente, la cuarta y última fase de la presente metodología prevé la aplicación de los indicadores fraseodidácticos a la fraseología extraída y anteriormente comparada con el corpus de lengua general del italiano. En este sentido, se avanzó una propuesta de clasificación de las unidades fraseológicas en función de su grado de frecuencia en la lengua actual italiana.

A continuación, se define pormenorizadamente el desarrollo de cada fase que compone el modelo metodológico secuencial exploratorio (Creswell, 2003, 2007, 2009, 2014) con excepción de la fase uno, revisión documental, puesto que su análisis se ha detallado en el marco teórico y el estado de la cuestión del presente trabajo (véase cap. II, III y IV).

5.2. Segunda fase: Extracción y anotación fraseológica

En la presente tesis doctoral se entiende la fraseología como una disciplina en estrecha relación con la lexicología, puesto que trata también la lengua común y, por otro lado, con la terminología, ciencia que estudia «Concepts and their representations in special language»[192] (ISO 704, 2000:vi). El fraseólogo es el especialista en fraseología de igual modo que el terminólogo lo es en terminología y el lexicógrafo en lexicografía.

Entre las principales tareas de quien trabaja en fraseografía destaca la creación y compilación de archivos fraseológicos, herramientas

[192] «Los conceptos y sus representaciones en la lengua especial». Traducción propia.

principalmente digitales que permiten la consignación y organización de los datos recabados de una muestra para informar sobre sus características (Pavel y Nolet, 2001). Dichos archivos fraseológicos pueden ser monolingüe, bilingüe o multilingüe y su creación se basa en un método de consignación y registro de las unidades recabadas. En el presente trabajo se ha optado por recurrir a los estándares terminológicos que desde la mitad del siglo XX han caracterizado la investigación contemporánea en terminología.

Basándose en la ISO 704: 2000[193], Silvia Pavel y Diane Nolet (2001: 15) en su ya clásico *Manual de terminología* presentan los contenidos que una ficha terminológica completa debe presentar. A saber:

- La presencia de una definición de los conceptos expresados.
- El uso uniforme de los términos que designan los conceptos.
- El uso limitado de variantes estilísticas, ortográficas y sintácticas.
- La medida en que trata los términos normalizados en el campo temático en cuestión.
- La justificación del uso o creación de nuevos términos.

Asimismo, el archivo fraseológico electrónico que se ha elaborado se basa en el principio uninocional (Pavel y Nolet, 2001: 23) según el cual a cada concepto corresponde una entrada. El principio uninocional, *modus operandi* empleado en las principales bases terminológicas mundiales[194], prevé que a una entrada corresponda un solo concepto. En este sentido, cada unidad representa un concepto monosémico, es decir, el exacto contrario del principio de polisemia que rige en los diccionarios generales y en el que una misma entrada corresponde una serie de acepciones que definen distintos conceptos. Las herramientas

[193] International Standard Organization (ISO) 704: 2000, Terminology work. Principle and methods.

[194] Entre otros, véase IATE, la mayor base de datos terminológica del mundo en la actualidad; EuroTermBank, la base de datos europea que centraliza la terminología bancaria para las lenguas de los nuevos estados miembros; ECHA-term, la base de datos terminológicas de la European Chemicals Agency; etc.

informáticas desarrollan un papel fundamental en cuanto permiten agilizar tanto la extracción (extracción automática, véase apartado 5.2.2.) de las unidades como su anotación.

A lo largo de esta sección se razona sobre las diferentes fases que han permitido elaborar el archivo fraseológico electrónico del presente estudio y sobre los métodos empleados en cada fase. En este sentido, se describe la elección motivada de un corpus representativo, el método empleado para identificar y extraer el contenido fraseológico y, finalmente, la anotación de la fraseología en un archivo fraseológico en formato digital elaborado *ad hoc* para el presente estudio. A continuación, se presenta y motiva la elección del corpus representativo escogido para que extraer la fraseología objeto de análisis.

5.2.1. Elección de un corpus representativo

La elección de un corpus representativo desde un punto de vista cualitativo y cuantitativo es fundamental para obtener resultados consistentes (Boas, 2010; Corpas y Seghiri, 2007). En el presente estudio se ha compilado un corpus a partir de dos series completas (del nivel A1 al C2) de dos manuales de referencia para la didáctica del italiano como lengua extranjera (ILE). La decisión de extraer fraseología a partir dos series completas de manuales no es baladí. En primer lugar, los manuales de una lengua extranjera, a menudo definidos como métodos[195], representan el telón de fondo en los cursos de idioma. A pesar de que estos no representen la única herramienta didáctica, el empleo de un manual es sin duda una práctica generalizada en las clases de LE.

En este sentido, los contenidos didácticos que conforman el manual representan, en línea general, el esqueleto sobre el que se dispone el programa didáctico de un curso de LE (Escribano, 2019). A ellos, el profesorado va alternando/integrando actividades externas al manual y

[195] Para más información véase <https://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/diccionario/metodo.htm>.

elaboradas en función de criterios que definen el alumnado como el nivel de conocimiento de la lengua, el número de asistentes, la duración de la actividad, la posibilidad de disponer de acceso a internet, etc.

Tomando dichas aseveraciones como punto de partida, se pueden proponer los siguientes dos axiomas: en primer lugar, los contenidos didácticos que el alumnado aprende pueden variar en función de la elección del manual y, por otra parte, el manual elegido representa el input primario para el alumnado. En este sentido, la elección de los manuales de una LE tendrá que ser guiada por diferentes criterios como la fecha de publicación, es fundamental que tanto los temas didácticos tratados como el léxico empleado sean actualizados; la editorial, que sea una referencia en la didáctica de la lengua objeto de estudio; los autores/as, con experiencia reconocida en la elaboración de materiales didácticos; la adaptación de los contenidos didácticos al nivel del alumnado; etc.

El corpus que se ha compilado para este trabajo doctoral consta de dos series completas de manuales de ILE; cada serie comprende seis manuales (uno por cada nivel del MCER) por un total de 12 manuales:

- *Nuovo Espresso*, Alma edizioni; niveles A1, A2, B1, B2, C1, C2.
- *Il Nuovo Affresco Italiano*, Mondadori Education; niveles A1, A2, B1, B2, C1, C2.

La representatividad del corpus es justificada por indicadores cualitativos y cuantitativos. Antes de adentrarnos en las características de cada uno ellos, y con el fin de definir los principales nexos entre ambos métodos, cabe destacar las características comunes que han llevado a la elección de dichos manuales. En este sentido, el presente corpus es representativo para el presente estdio en cuanto se trata de manuales:

- De reconocido prestigio y ampliamente utilizados en enseñanza de italiano como lengua extranjera: ambos manuales se emplean en

diferentes sedes de la Sociedad Dante Alighieri, y han sido utilizados también en la Escuela Oficial de idiomas de Málaga.

- Completos: cada editorial dedica un manual específico a cada uno de los niveles del MCER (A1-C2) a través de un programa anual.
- Consistentes: tratan con la misma metodología los seis niveles del MCER (A1-C2).
- Recientes: con el fin de extraer fraseología actual, su primera edición es posterior a 2014.
- Fiables: tanto las editoriales, como los autores y autoras que escribieron los manuales, son considerados referentes en la didáctica del italiano como lengua extranjera.

A continuación, se detallan pormenorizadamente las dos series de manuales de ILE elegidos para compilar el corpus y, posteriormente, extraer la fraseología relevante para el presente estudio.

5.2.1.1. Nuovo Espresso, Alma Edizioni

Alma Edizioni, es líder mundial en la producción de materiales para la didáctica del ILE. Fundada en 1994, en más de 25 años supo imponerse en el mercado internacional del italiano como lengua extranjera vendiendo en más de 80 países[196]. En su catálogo ofrece más de 150 títulos entre cursos de idiomas y materiales complementarios destinados a un amplio target que va de niños a adultos. Asimismo, Alma Edizioni es partner oficial de la Società Dante Alighieri[197] y, según su página web, el 35% de los *Istituti Italiani di Cultura* (ICC)[198] emplean esta editorial en sus clases de ILE.

[196] Para más informaciones véase: https://www.almaedizioni.it/it/informazioni/chi-siamo/.

[197] Fundada en 1889, la sociedad Dante Alighieri a través de casi 500 comités en todo el mundo protege y difunde la lengua y la cultura italiana a través de academias, bibliotecas, cursos de lengua italiana, conferencias y manifestaciones artísticas de diferentes niveles y tipologías. Al día de hoy, la Dante Alighieri es la mayor red de centros de formación de ILE en el mundo.

[198] Dependientes del Ministerio de Asuntos Exteriores italiano, los *Istituti Italiani di Cultura* suman un total de 83 centros por el mundo y sus finalidades son la difusión de la lengua y cultura italiana en el mundo.

Entre sus productos estrella se encuentra el método *Nuovo Espresso*, un curso de italiano como lengua extranjera dividido en seis niveles (A1, A2, B1, B2, C1, C2) en línea con las indicaciones del MCER. Cada manual aborda un nivel independiente del MCER y, aunque el método empleado sea el mismo, diferentes autores han contribuido a su redacción como se detalla a continuación en la Tabla 1:

Tabla 1. Autores de los manuales *Nuovo Espresso* por nivel.

Nivel	**Autores**	**Horas**
A1	Luciana Ziglio y Giovanna Rizzo	90
A2	Maria Balì y Giovanna Rizzo	90
B1	Maria Balì y Luciana Ziglio	90
B2	Maria Balì e Irene Dei	90
C1	Giorgio Massei y Rosella Bellagamba	90
C2	Michela Guida y Chiara Pegoraro	90

Fuente: elaboración propia.

El manual es la continuación natural de *Espresso*[199], cuya metodología y contenidos se habían convertido en obsoletos y necesitaban una actualización significativa para seguir siendo referentes en el mercado de ILE. Con *Nuovo Espresso*, la editorial no se ha limitado a modernizar los contenidos didácticos, sino que ha creado un nuevo método, proponiendo nuevos contenidos como lecturas, ejercicios ampliados y actualizados, un videocurso e incluso una videogramática para profundizar en aspectos lingüísticos, estructuras y funciones comunicativas. Cabe destacar que se echa mucho en falta una presentación o introducción general y detallada al método empleado.

[199] Según su página web vendió más de un millón de copias en todo el mundo (véase: https://www.almaedizioni.it/it/informazioni/chi-siamo/).

En este sentido, solo los niveles A1, A2 y B1 ofrecen una breve introducción en la que se presenta el manual.

Cada nivel de *Nuovo Espresso* se compone por diez unidades didácticas llamadas *lezioni* (clases). Cada unidad ofrece ejercicios para practicar las cuatro destrezas: producción oral y escrita, comprensión lectora y auditiva. Al final de cada unidad, el manual ofrece una sección en la que se presenta un resumen de las competencias gramaticales estudiadas a lo largo de la unidad y, aproximadamente cada cinco unidades, propone una prueba global en el que se recopilan los contenidos de las unidades anteriores. Asimismo, en la parte final de los manuales, se ofrece al alumnado una sección de ejercicios gramaticales para practicar en casa y repasar las competencias aprendidas en el aula.

Desde un prisma fraseodidáctico, se ha analizado si el manual manifiesta una intención explícita con respecto al tratamiento de la fraseología en uno de sus apartados (introducción general, introducción de cada unidad, gramática, etc.) y, de ser así, cómo se trata a lo largo del método. En este sentido, en la brevísima introducción al manual no se hace ninguna mención al tratamiento que la fraseología recibe en el manual. Asimismo, el análisis de los índices de los manuales de los distintos niveles manifiesta la misma ausencia previamente descrita con respecto al tratamiento de las UF.

Sin embargo, cada nueva unidad didáctica comienza con una página introductoria en la que se presenta una pincelada general de las funciones comunicativas, la gramática y el léxico que se aprenderá a lo largo de la unidad. En dichas páginas introductorias, se observa que el manual proporciona muestras de la fraseología que se aprenderá en la unidad bajo el apartado dedicado al léxico. Sin embargo, se observa cierta confusión terminológica a la hora de identificar la terminología empleada para definir las UF. En el mismo método se emplean distintos términos para definir el mismo objeto de estudio: *Espressioni idiomatiche*

(unidad 1, nivel C1 | unidad 5, nivel C2), *Espressioni fisse* (unidad 3, nivel C1|unidad 2, nivel C2), o simplemente *Espressioni* (unidad 7, nivel C1).

A pesar de las limitaciones encontradas en los apartados introductorios (presentación del manual, índice y páginas introductorias de cada unidad) al adentrarse en las diez unidades didácticas de cada nivel se observa cierto interés en la fraseología y, en algunos casos, se le dedica una sección de ejercicios exclusiva. Sin embargo, la confusión terminológica descrita previamente se repite y enfatiza a lo largo de los manuales. La primera mención a la fraseología se observa ya a partir del nivel A1 (p. 88) y, en particular, se presenta una actividad sobre las locuciones, que denomina *Modi di dire* sin ofrecer una definición del término empleado. Asimismo, en el nivel C1 (p. 29), dicha esfera de las locuciones es llamada *Espressioni fisse*, mientras que en un audio de la siguiente unidad (pista 6) se presenta un ejercicio sobre las *Frasi fatte* que sucesivamente define como *Espressioni idiomatiche*.

Este último término es el mismo que se emplea unas páginas más adelante (nivel CI, p. 112) pero, lo que realmente propone el ejercicio es completar una colocación con el colocativo adecuado. En el mismo nivel C1 (p. 106) el manual propone un ejercicio sobre las paremias que en un primer momento llama *Proverbi* y sucesivamente *Modi di dire*, empleando los dos términos como sinónimos del mismo concepto. Finalmente, una página después emplea el término *Segnali discorsivi* (p. 107) para identificar las fórmulas, aunque solo tres páginas antes los identificaba como *Connettori ipotetici* (p. 104).

En conclusión, se puede afirmar que por un lado no se proporciona una definición de los conceptos empleados y, por otra parte, no se toman en consideración las propuestas de clasificación presentadas en las últimas décadas por la academia italiana y que han sido resumidas en el marco teórico del presente estudio[200]. Se ha observado que los manuales no

[200] Véase capítulos II y III.

consiguen dar consistencia terminológica a los conceptos ocasionando confusión tanto en el profesorado como en el alumnado. Los términos empleados para definir las clases o esferas de la fraseología no son consistentes ni armonizados y, a menudo, se solapan o se utilizan como sinónimos para describir conceptos diferentes entre ellos.

Sin embargo, y a pesar de dichas limitaciones terminológicas, se puede afirmar que las esferas que componen el universo fraseológico son tratadas ampliamente en los manuales de *Nuovo Espresso* a través de ejercicios exclusivos para su descifración y memorización y el método analizado presenta un número significativo de unidades fraseológicas. En conclusión, se puede afirmar que dichos manuales presentan una amplia gama de UF aunque falle en la clasificación de las mismas.

5.2.1.2. Il Nuovo Affresco Italiano, Mondadori Education

Mondadori Education nace en 1988 como editorial del Gruppo Mondadori dedicada a la educación y formación a través de dos áreas: por un lado, escolástica y diccionarios, área enfocada hacia la escuela primaria y secundaria, e italiano como lengua extranjera y universidad, por otro lado. Asimismo, Mondadori Education es ente acreditado por el Ministero dell'Istruzione, dell'Università e della Ricerca[201] (MIUR) para la formación del personal escolástico.

Con respecto a la didáctica del italiano como lengua extranjera, actualmente su producto estrella es *Il Nuovo Affresco Italiano*. Dicho manual de ILE recoge la herencia de *Affresco Italiano* cuya primera edición fue publicada en 2007. La nueva edición es actualizada con respecto a los contenidos didácticos y ampliada en cuanto a número de páginas con respecto a su predecesor. El curso, cuyos autores son Maurizio Trifone y Andreina Sgaglione, se articula en los seis niveles (A1,

[201] Homólogo italiano del Ministerio de Educación, Universidades e Investigación.

A2, B1, B2, C1, C2)[202] del MCER y cada uno de ellos se presenta en un volumen único que comprende el manual para trabajar en el aula y la recopilación de ejercicios al final del libro, ambos estrechamente ligados entre ellos en cuanto a la estructura y progresión de los contenidos.

Los manuales se desarrollan a través de doce unidades didácticas para los niveles A1, A2 y B1, nueve unidades para el B1 y B2 y cinco unidades para el C1 y C2. Asimismo, las unidades didácticas se dividen a su vez en ocho subsecciones llamadas *lezioni* (clases) que incluyen una introducción al tema tratado en la unidad, una sección dedicada a la gramática, una al léxico que se aborda en la unidad, una serie de ejercicios para practicar las cuatro destrezas y, finalmente, consideraciones y ejercicios fonético. Al final de cada unidad se ofrece también un corolario gramatical, en el que se resumen los principales aspectos gramaticales tratados en el aula, y un test de una página para poner a prueba el alumnado sobre los contenidos gramaticales aprendidos. Un aspecto interesante de dicho manual es su digitalización. En este sentido, la editorial pone a disposición del profesorado y del alumnado il *MEbook*, un manual digital rico en contenidos, videos, audios y ejercicios. Un conjunto muy valioso que complementa el manual a través de materiales didácticos integrados.

Analizando los manuales desde un prisma fraseodidáctico, se observa que en su introducción no se hace referencia explícita a la fraseología y a como es tratada a lo largo de curso. En cambio, estudiando los índices de los manuales se observa que se hace primera mención explícita a la fraseología a partir del nivel B2 (unidad 2) y, en particular, a través de un ejercios denominado *Proverbi e espressioni idiomatiche*. Por otra parte, la última sección de los manuales es dedicada a un glosario bilingüe[203]

[202] Cabe destacar los niveles C1 y C2 se encuentran en elaboración por lo que el análisis y la extracción fraseológica de dichos niveles se realizó en *Affresco Italiano* (C1 edición 2019 y C2 edición 2018) cuyos autores son Maurizio Trifone, Andreina Scaglione y Antonella Filippone.

[203] En el MEbook se encuentra una versión plurilingüe en la que además del inglés se puede encontrar en español, alemán, portugués, francés y ruso.

(italiano-inglés) en el que se presenta el léxico y la fraseología aprendida en el nivel según la siguiente estructura:

palabra italiana [transcripción fonética] *categoría gramatical* • traducción/es al inglés (número de la unidad en el que compare el vocablo por primera vez)
Ej. **bosco** [bosko] *nome* • wood (11)
Ej. **acqua gassata** [akkwa gas'sata] • carbonated water, sparkling water (9)
Como se observa en los ejemplos proporcionados, el glosario resulta particularmente interesante en cuanto ofrece también fraseología ordenada según criterio alfabético. Cuando se menciona una unidad fraseológica, el campo que expresa la categoría gramatical es omitido, lo que no ocurre en el léxico simple. Dicha elección que atañe exclusivamente a la fraseología es significativa porque denota cierta dificultad de los autores del manual a la hora de encasillar la fraseología en categorías gramaticales definidas.

A pesar de dicha limitación, el hecho de que los autores hayan decidido incluir la fraseología en el glosario es llamativo. En este sentido, el hecho de poner bajo un mismo contenedor el léxico simple y la fraseología es significativo dado que, como se ha visto en el marco teórico (cap. II), la fraseología en Italia se ha tratado y se sigue tratando desde una perspectiva lexicográfica. Cabe destacar que la inclusión de la fraseología en el glosario se observa solo en los niveles A1 y A2. A partir del B1 desaparece y los distintos glosarios presentan exclusivamente léxico simple, lo que supone un empobrecimiento del contenido fraseodidáctico.

En cambio, a partir del nivel B1 (unidad 4, p. 78) empiezan a proponerse ejercicios específicos sobre el uso de la fraseología. En ellos, se abordan las esferas de las locuciones y las fórmulas rutinarias, aunque los autores las definen bajo el término genérico de *Espressioni*. Una vez más, dicho

manual reafirma la tendencia de proponer ejercicios sobre la fraseología a partir de los niveles intermedios, omitiendo los más básicos.

Con respecto al uso de la terminología para clasificar las diferentes unidades fraseológicas, una vez más, se observa un alto grado de inconsistencia. A lo largo de los diferentes manuales, las UF son definidas como: *Espressioni* (unidad 4, p. 78, nivel B1), *Espressioni idiomatiche* (unidad 2, p. 40, nivel B2), *Proverbi e espressioni idiomatiche* (unidad 2, p. 40, nivel B2), *Locuzioni verbali* (unidad 2, p. 56, nivel C1), *Espressioni fisse* (unidad 3, p. 76, nivel C2), *Espressioni burocratiche* (unidad 2, p. 43, nivel C1), etc. En este sentido, por un lado se observa una ausencia de armonización con respecto a los términos que describen el concepto y, por otro lado, una ausencia completa de definiciones que identifiquen los conceptos.

A partir del nivel B2 se introduce la distinción entre dos conceptos estrechamente ligados con la idiomaticidad: significado literal frente al significado figurado. No se propone una definición clara de dichos conceptos sino se ofrecen ejercicios en el que el alumnado tiene que interpretar de manera deductiva el correcto significado de cinco locuciones escogiendo entre tres posibilidades (unidad 1, p. 24, nivel B2). Tras su primera aparición en el nivel B2, el manual propone en todas las unidades sucesivas un breve ejercicio sobre los *Proverbi* con el objetivo de hacer reflexionar sobre la idiomaticidad. El método didáctico se repite: se le pide al alumnado que deduzca y explique en el aula el significado de dos o tres refranes acompañados por sus respectivas ilustraciones. La presencia de un apartado específico de refranes en el nivel B2 deja patente la intención de los autores que consideran dicho nivel como el más adecuado para enseñar la paremiología.

En conclusión, a pesar de que los términos empleados para definir las esferas del universo fraseológico no sean consistentes (*espressioni, espressioni idiomatiche, espressioni fisse, locuzioni verbali*, etc.) y que no

se ofrecen definiciones para aclarar los conceptos, se puede afirmar que los manuales del *Il Nuovo Affresco Italiano* tratan la fraseología como parte integrante de las unidades didácticas. En particular, dicho tratamiento aparece tímidamente a partir del nivel B1 y con más fuerza a partir del nivel B2 en el que no solo se introduce el concepto de idiomaticidad, sino se trata con ejercicios *ad hoc* la esfera de las locuciones y de las paremias. Asimismo, en el índice del nivel B2 se menciona la fraseología demostrando cierta sensibilidad hacia la fraseodidáctica.

Por otra parte, al final de los manuales A1 y A2 se ofrece un amplio glosario en el que, además del léxico simple, figuran colocaciones, locuciones y fórmulas rutinarias. Dicha presencia de fraseologismos en los glosarios pone la fraseología y el léxico en un plan equiparable, y el hecho de que la fraseología desaparezca en los glosarios de los siguientes niveles representa sin duda un empobrecimiento de los contenidos fraseodidácticos.

5.2.2. Extracción y anotación de las UF

La fraseología computacional, en su término inglés *computational phraseology* (Rothkegel, 2008; Heid, 2007, 2008; Barrón-Cedeño, 2019; Gupta, 2019; Corpas y Colson, 2021, Corpas, 2022), se suele definir como la aplicación de las tecnologías lingüísticas para detectar, extraer, estudiar y representar las unidades fraseológica pertenecientes a uno o más idiomas. Sus principales aplicaciones prácticas o áreas de estudios vienen a ser dos (Heid, 2007, 2008; Corpas Pastor, 2013): por un lado, los métodos y herramientas informáticas para procesar textos que presentan UF. En este sentido, los métodos de traducción automática (TA) mediante sistemas neuronales podrían representar un ejemplo actual. Por otra parte, la herramientas y aplicaciones para extraer fraseología y/o para localizar fraseología ilustrativa y estudiar su uso en la lengua escrita o

hablada. Esta última aplicación representa el objeto de estudio del presente capítulo.

La extracción terminológica es una de las principales tareas de quien realiza estudios lingüísticos de corpus especializados con diferentes objetivos como la creación de diccionarios, elaboración de thesaurus, clasificaciones, memorias de traducción (MT), base de datos terminológicas (TB), aprendizaje ontológico, etc. La extracción puede ser monolingüe o bilingüe. En el primer caso, se dirige al análisis de textos o de un corpus compuesto por un solo idioma, mientras que en la bilingüe se compara el texto de origen con el texto meta para identificar candidatos a términos y sus equivalentes.

En el presente estudio doctoral se ha recurrido principalmente a las normas ISO en materia de terminología para proceder a la extracción de la fraseología a partir de un corpus *ad hoc* previamente compilado y descrito anteriormente. Extraer las unidades fraseológicas relevantes a partir de un corpus dado es un trabajo largo que absorbe mucho tiempo y requiere paciencia. Dicha labor se puede realizar principalmente de dos formas: extracción automática o manual.

En las últimas décadas se han desarrollado programas informáticos que reconocen y extraen de manera automática unidades terminológicas y fraseologicas (Plum, 2022; Hidalgo-Ternera, 2020)); en otras palabras, se trata de herramientas informatizadas creadas con el fin de ayudar en el trabajo de vaciado, permitiendo la extracción automática de candidatos a términos a partir de un corpus dado (Cruz Diaz, 2015). Cabe destacar que, a pesar de que dichas valiosas herramientas ayudan y facilitan la extracción, el resultado tiene que pasar por un filtro humano puesto que dicho tipo de extracción puede generar por un lado una gran cantidad de ruido y, por el otro, silencio[204].

[204] Con el término ruido se entienden «expresiones seudoterminológicas o elementos que aparecen juntos en el discurso accidentalmente pero que no designan conceptos particulares» (Pavel y Nolet

En palabras de Corpas Pastor (2013: 342), «La identificación de unidades fraseológicas en los corpus constituye una de las dificultades por antonomasia de la fraseología computacional». Actualmente, la extracción automática de términos (ATE, en su sigla del inglés *Automatic terminology extraction*) se basa principalmente en tres métodos[205] (Vázquez y, Oliver, 2018): lingüísticos, estadísticos e híbridos. El enfoque lingüístico permite extraer términos a partir de determinados patrones sintácticos o morfológicos previamente establecidos y combinados con un corpus etiquetado en función de las partes del discurso como, por ejemplo, las combinaciones verbo+nombre, nombre+nombre, etc. La limitación, y fuerza al mismo tiempo, de este enfoque reside en la fuerte dependencia con la estructura de lengua objeto de análisis. En este sentido, por cada idioma se deben emplear patrones diferentes en función de su estructura sintáctica, lo que conlleva que dicha metodología sea empleada sobre todo en extracción monolingüe o entre lenguas muy afines.

El enfoque estadístico, en cambio, se basa en las secuencias repetidas de palabras. Se trata de un enfoque que gravita en la independencia de la lengua, y encuentra su fuerza en la estadística y la frecuencia. En este sentido, la frecuencia de aparición de un término o de una unidad en un corpus asume un papel prioritario y, por ello, numerosas herramientas que utilizan dicho enfoque permiten al usuario establecer un número mínimo y máximo de repeticiones para considerar la extracción como significativa.

Finalmente, el enfoque híbrido elabora sus algoritmos tanto en función de la información estadística como de la lingüística en un corpus dado. Dicha aproximación representa actualmente la metodología más

2001:48). En cambio, con silencio se entiende la ausencia de expresiones terminológicas presentes en el corpus que la herramienta de extracción no identifica.

[205] Cabe destacar que nuevas propuestas se han llevado a la mesa del debate en los últimos años. Véase pro ejemplo, las interesantes propuestas de Rohanian et al. (2019), y de Wahl y Gries (2018), etre otros.

empleada[206] en las herramientas de extracción puesto que permite extraer estadísticamente la terminología teniendo en cuenta los criterios lingüísticos, lo que permite afinar la búsqueda y eliminar buena parte de ruido. Las principales herramientas de pago que ofrece el mercado para el procesamiento de expresiones multiverbales[207] son SynchroTerm de Terminotix, MultiTerm Extract de RWS Group (antes de principio de 2021 era propiedad de SDL Trados) y Sketch Engine de Lexical Computing Limited.

SynchroTerm, se basa en un enfoque estadístico y emplea sus algoritmos para extraer términos de un corpus monolingüe o bilingüe. Se trata de una herramienta muy valiosa en cuanto permite elaborar archivos terminológicos de manera intuitiva y eficaz para exportarlos en diferentes formatos compatibles con los principales CAT tools: Excel, HTML, SDLTB (SDL Multiterm), Trados WinAlign, etc. Asimismo, soporta 30 idiomas y permite afinar y personalizar los criterios de búsqueda de forma personalizada y bastante avanzada. El principal problema que presenta dicha herramienta tiene que ver con el ruido al que la extracción da lugar: palabras comunes en lugar de términos, secuencias de palabras no relevantes, etc., que complican el análisis humano a posteriori.

Por otra parte, a través de algoritmos basados en la frecuencia, SDL MultiTerm Extract de SDL Trados (RWS Group) permite extraer automáticamente términos de un corpus para elaborar glosarios, bases de datos e incluso compilar diccionarios a partir de textos en paralelo en dos idiomas. Sin embargo, esta aplicación es poco intuitiva y los parámetros para definir la extracción necesitan ser afinados con esmero so pena la pérdida de información valiosa.

---

[206] Véase la Terminology Coordination Unit of the European Parliament (TermCoord): <https://termcoord.wordpress.com/2013/07/05/why-terminology-extraction/>.

[207] Del inglés «multi-word expresions (MWE)», dicho término más frecuente en os trabajos de PNL, es equiparable al de unidades fraseológicas (Corpas Pastor, 2013). Sin embargo, y dado el enfoque fraseodidáctico de la presente investigación, se prefiere utilizar el término unidad fraseológica.

Finalmente, Sketch Engine es los que se define como un *Corpus Query Systems* (CQS) que, entre sus funcionalidades, permite extraer terminología o fraseología a partir de amplios corpus. Actualmente esta herramienta es más completa con respecto a las que se han presentado anteriormente puesto que, a través de una única interfaz, permite acceder a numerosas funciones entre las que destaca el análisis de concordancia por términos, frases o unidades; la comparación de un corpus propio con uno de referencia para estudiar los términos extraídos en corpora de gran tamaño; la extracción de términos a partir de un enlace web sin la necesidad de descargar sus archivos y compilar un corpus manualmente; el estudio de la frecuencia de unidades simples en corpora de grande tamaño, etc.

Con respecto a las herramientas de extracción gratuitas, se menciona TermoStat Web 3.0 y AntConc. TermoStat Web 3.0 es gratuíto, pero es necesario darse de alta en la página web. Emplea el método híbrido para extraer términos a partir de diferentes formatos: listas, glosarios, corpus, etc. Por otra parte, AntConc es una aplicación desktop descargable por Windows, Macintosh OS X, y Linux que, a partir de un corpus propio, permite analizar concordancias, n-gramas, colocativos, word list, etc. Dichas herramientas son customizables en función de las necesidades del investigador y permiten proporcionar válidos candidatos en función de la búsqueda que se realiza.

Por último, se menciona el *Term Extracción Module* (TEM), la reciente[208] herramienta de extracción terminológica presentada en la versión 2.13.1 por el IATE Support & Development Team (Interactive Terminology for Europe) y en continuo desarrollo (véase figura 6). El TEM permite extraer unidades simples o unidades fraseológicas que aparecen con más frecuencia en un corpus dado que en la lengua general. La extracción se realiza comparando el corpus propio con un *chuncker* de referencia

[208] Presentada el 17 de febrero de 2020 y en continua evolución. Actualmente es disponible solo para socios internos de IATE por lo que su utilizo necesita ciertos permisos a nivel de usuario.

(TreeTagger[209]) y de momento es disponible para inglés, español, alemán e italiano[210].

Los resultados de la extracción son presentados en dos tablas distintas, por un lado las unidades simples y, por el otro, las unidades fraseológicas. Cada candidato a término dispone de una serie de metadatos (frecuencia absoluta, frecuencia relativa, etc.) y se puede estudiar a través de un máximo de 20 líneas de concordancia (*occurrencies)*, es decir, ejemplos de aparición de un término en su contexto original. Cabe destacar que para acceder a dicha herramienta es necesario disponer, en primer lugar, de una cuenta ECAS[211] y, en segundo lugar, tener permisos específicos[212] que permitan trabajar con el TEM. Con esta herramienta, IATE supera su vocación original de base terminológica de la Unión Europea y se presenta como herramienta holística con mirada hacia las herramientas de traducción asistida por ordenador (TAO).

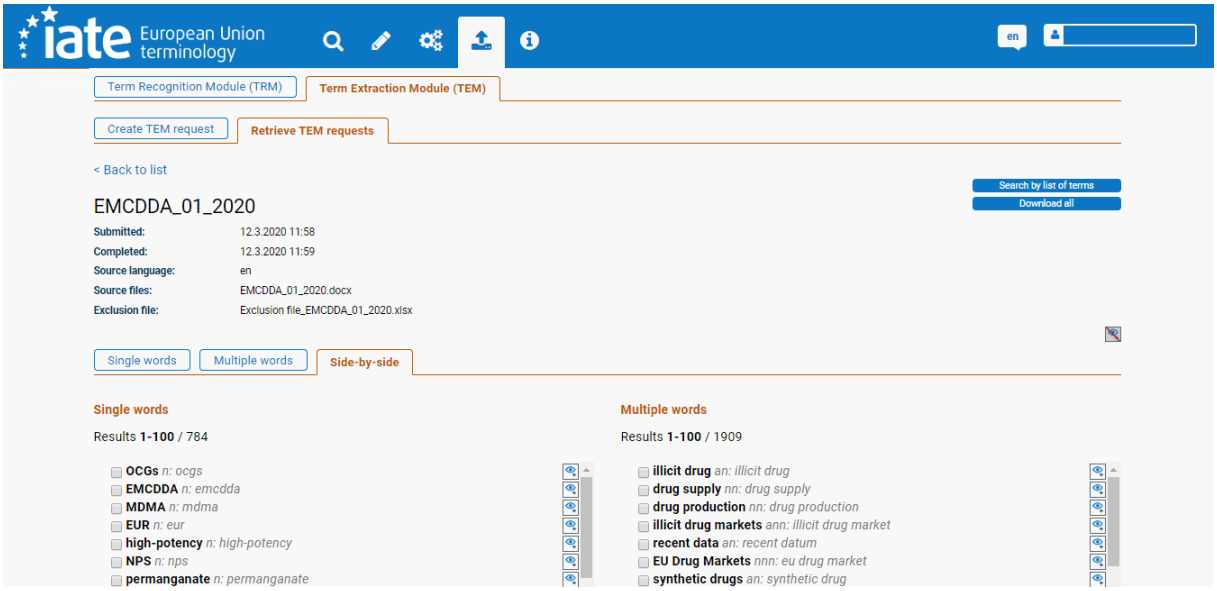
[209] Herramienta elaborada por Helmut Schmid en el Institute for Computational Linguistics de la Universidad de Stuttgart para anotar texto en función del lema o del elemento sintáctico que desarrolla en el texto. Para más información sobre el pre y post procesamiento del corpus véase el *Handbook* de IATE: https://iate.europa.eu/assets/handbook.pdf.

[210] Para otros idiomas el TEM emplea el inglés como modelo de comparación, lo que supone cierta limitación en los resultados esperados.

[211] Sistema de autenticación de la Comisión Europea (The European Commission Authentication System en inglés). Se trata de un servicio gestionado por la Comisión que permite el acceso a los servicios digitales utilizados por las Instituciones Europeas.

[212] Los utentes con el rol *basic user* no tienen los permisos para acceder al Term Extraction Module.

Figura 7. Term Extraction Module (TEM) de IATE.



Fuente: https://iate.europa.eu/home.

Todas las herramientas de extracción automáticas descritas en este epígrafe son de gran utilidad no solo para la investigación, sino también para terminólogos y traductores profesionales que podrán emplearlas para compilar o mejorar sus bases de datos (TB) y aplicarlas a una traducción o a una memoria de traducción (MT). Para emplear una de estas herramientas, en necesario que el archivo que se pretende analizar sea en formato electrónico y, normalmente, todas estas herramientas de extracción reconocen las extensiones más comunes como .DOCX, .XLSX, .PPTX, .TXT, .PDF[213], etc., pero también. TMX u otros formatos de marcado como .XML, .HTML, .CVS, etc.

Como se definirá pormenorizadamente el apartado 5.2.1. del presente capítulo, en la presente investigación se optó por realizar una extracción manual de la fraseología al no disponer del formato electrónico de los contenidos que se pretendía analizar, sino solo de su versión impresa. Sin embargo, se realizó un pretest con el objetivo de averiguar si, convirtiendo al formato digital los manuales que componen el corpus, se

[213] Se aconseja evitar dicho formato al no ser siempre compatible al cien por cien con las herramientas de extracción automática.

podía contar con algunas de las herramientas de extracción automática previamente descritas.

Por todo ello, la extracción fraseológica llevada a cabo en el presente estudio doctoral ha sido realizada a través de tres fases consecutivas que se definen a continuación:

1. Pretest para la extracción automática de las unidades a través de tres herramientas de extracción: SynchroTerm, Sketch Engine y el Term Extraction Module de IATE.

2. Identificación y extracción manual de las unidades fraseológicas a partir dos series completas de manuales de italiano L2.

3. Anotación y transferencia de fraseología extraída en el archivo fraseológico monolingüe en función del principio uninocional.

En el siguiente apartado se describe el pretest, así como sus resultados.

5.2.2.1. Pretest: extracción automática con SynchroTerm y Sketch Engine

Extraer manualmente fraseología a partir de un corpus compuesto por dos series completas (A1 - C2), es decir, 12 manuales de ILE y más de 3.600 páginas, requiere tiempo y esfuerzos significativos. Con el objetivo de agilizar la extracción, se ha optado por realizar un pretest para averiguar la eficacia de las herramientas descritas en el apartado anterior. Para ello se ha procedido de la siguiente manera.

En primer lugar, al no disponer de los manuales en un formato compatible con las herramientas de extracción, se ha escaneado manualmente una muestra de diez páginas a partir de uno de los manuales objeto de estudios[214]. Dicha fase previa ha supuesto *per se* cierta inversión de tiempo que hay que tener en consideración. El resultado de dicha fase

[214] Nuovo Espresso Italiano, nivel C2, de la página 162 a la 173 del manual. Se ha elegido dicho nivel por suponer que en este nivel se pudiese encontrar una mayor cantidad de UF pertinentes para el presente estudio.

ha sido un .PDF no editable, es decir, similar a un .JPG, que ninguna herramienta de extracción reconoce como input. Para obviar al problema, se ha empleado el plan Acrobat Pro DC[215] de Adobe para convertir .PDF no editables a un formato admitido por las herramientas de extracción, en este caso, un nuevo documento de Microsoft Word con extensión .DOCX.

Para testar la extracción se recurrió a Sketch Engine siendo esta una de las herramientas más valiosas para extraer términos o n-gramas de un corpus dado. Al introducir el corpus en la herramienta, se ha impostado los parámetros de Sketch Engine para que reconociera entre dos y cinco n-gramas con una frecuencia mínima de aparición en el corpus de uno. Sin embargo, y como se explica a continuación, los resultados no fueron satisfactorios. La herramienta reconoció 287 unidades de las cuales solo 8 representaban unidades fraseológicas[216] relevantes. Como se puede observar, el ruido causado por la herramienta es exponencialmente superior a los resultados deseables. Puesto que los resultados no fueron consistentes, se realizaron otras pruebas para mejorar la extracción introduciendo una *stop list* que incluía, entre otros, artículos, preposiciones y pronombres. De este modo, los resultados mejoraron ligeramente devolviendo 305 unidades de las cuales 11 eran unidades fraseológicas[217] reales.

Tras diferentes pruebas inconsistentes en las que se modificaron los parámetros de búsqueda de Sketch Engine, se optó por emplear otra herramienta de extracción terminológica, SynchroTerm, utilizando los mismos parámetros empleados en el test anterior. En este caso, la herramienta reconoció 1310 unidades de las cuales solo 13 unidades

[215] Plan completo y profesional que permite acceder a todas las herramientas disponibles de Adobe.
[216] A saber: *essere terra terra, stare quatto quatto, quasi quasi, caffé ristretto, al posto giusto, passo passo, parere contrario a, produzione cinematografica.*
[217] Las anteriores más *anche se*, al *posto giusto, scosse di terremoto.*

fraseológicas[218]. A pesar de que dicha herramienta haya conseguido extraer un número mínimamente superior al caso anterior, el ruido causado fue más de cuatro veces superior al test precedente, lo que en meros términos de tiempo significó mucho más esfuerzo para identificar las unidades reales.

Finalmente, se optó por comparar los resultados obtenidos a través de las dos extracciones automáticas con una extracción manual. En este sentido, la extracción manual permitió identificar 36 UF, es decir el 292,3% más de unidades encontradas manualmente frente a la extracción realizada a través de SynchroTerm y hasta el 345,4% más que Sketch Engine.

Los resultados deficitarios obtenidos a través de la extracción automática pueden ser atribuidos a dos causas principales. En primer lugar, el tipo de formato del corpus introducido en las herramientas no era el ideal. Al convertir un documento .pdf no editable en uno editable se pierde parte de la formatación del texto: separación de las palabras como, por ejemplo, ocurrió en la frase *Mia sorella maggiore si è sposata in Inghilterra* que durante la conversión fue modificada en *Mia sorella maggio re si è sposat a in Inghilte rra*. Los espacios gráficos son considerados por las herramientas de extracción como indicadores de fin y comienzo de otra palabra lo que causó una disfunción en el algoritmo. Está claro que solucionar el problema corrigiendo manualmente, o a través de REGEX[219], el texto conlleva un esfuerzo que, en términos de tiempo, no compensa el proceso de extracción.

En segundo lugar, se cree que tanto SynchroTerm como Sketch Engine proporcionen mejores resultados para la lengua inglesa y que en otros

[218] A saber: aggiustare un guasto, padrone di casa, a modo mio, senso dell'orientamento, per favore, sul da farsi, la febbre passa, fratturarsi una gamba, ciel sereno, in seguito a, uomo avvisato mezzo salvato, tamponamento a catena, caffè ristretto.

[219] El acrónimo proviene del inglés *Regular expression* y se ha traducido al español como Expresiones regulares.

idiomas sean aún mejorables. Finalmente, cabe destacar que la comparación entre las dos extracciones automáticas y la manual ha permitido observar dos interesantes aspectos: ambas herramientas, aunque SynchroTerm de manera más significativa, han devuelto por un lado un gran ruido, lo que significa tiempo y esfuerzo para el investigador que tenga que analizar manualmente los resultados de la extracción y, por otro lado, un gran silencio, es decir, pérdida significativa de información relevante para el estudio.

En conclusión, el principal objetivo de emplear una herramienta de extracción reside en agilizar la investigación en términos de tiempo y esfuerzo. Sin embargo, en este caso no solo los resultados fueron insatisfactorios, sino que se tuvo que dedicar tiempo y energías para escanear los manuales, convertir el formato, limpiar los errores de formato y analizar la extracción. Por todo ello, y a pesar del gran esfuerzo que esta decisión supuso, se optó por extraer manualmente las unidades fraseológicas de los manuales objeto de estudio con el objetivo de obtener mejores resultados desde un punto de vista cualitativo y cuantitativo.

5.2.2.2. Identificación y extracción manual de la fraseología

Si bien la decisión de proceder con la extracción manual de la fraseología, a pesar del gran tamaño del corpus, se vio obligada por las limitaciones ocasionadas por la informatización del proceso, es importante destacar las ventajas que una extracción de tipo manual ofrece frente a una automática. A pesar de los logros que en la última década las herramientas de extracción automática han mostrado, la extracción automática de la fraseología sigue siendo al día de hoy uno de los mayores retos de la lingüística computacional (Corpas Pastor y Colson, 2020) puesto que extraer fraseología de manera automática es un proceso más complejo que la extracción del léxico simple por varias razones entre las que destaca, por un lado, la relación semántica que se

establece entre los componentes unitarios de una unidad fraseológica (discontinuidad, colisión homonímica, etc.) y, por otro lado los errores relacionados con el proceso de etiquetado POS, parsing, ortográficos en el mismo corpus, substandard spellings, etc., que son una constante con la que lidiar, y limitar, durante el proceso de etracción (Gupta, 2019; Hidalgo Ternero, 2021; Corpas Pastor, 2021; Plum, 2022, etc.).

Los programas de gestión automática de los corpus lingüísticos desarrollan un trabajo fundamental para elaborar estudios cuantitativos, pero el investigador «debe cribar los datos de acuerdo con una serie de criterios, generalmente de tipo semántico» (Corpas Pastor, 2001: 100). Analizando, por ejemplo, el concepto de colocación, en el que prima la atracción semántica entre colocativo y colocado, Corpas Pastor observa que las herramientas de extracción pueden verse limitadas puesto que en un texto natural «hay colocaciones cuyos elementos aparecen muy distanciados en el discurso, por lo que no pueden ser extraídos de forma automática» (Corpas Pastor, 2001: 100).

Asimismo, la colisión homonímica (Penadés, 2015), es decir, la homonimia que se produce entre una locución y una combinación de palabras de índole no fraseológica es un fenómeno que solo un atento análisis humano puede descifrar. Desde esta perspectiva, frente a una secuencia como *quitarse el sombrero*, la máquina no puede reconocer si se trata de una locución idiomática o menos. En cambio, la evaluación humana realizada por un hablante nativo, o con un buen nivel de conocimiento, permitiría reconocerlo en seguida. En este sentido, el análisis humano prevalece desde un punto de vista cualitativo frente a un extractor automático. Sin embargo, la principal limitación del proceso manual frente al automático es el tiempo necesario para elaborar la extracción, siendo en el primer caso exponencialmente superior.

El corpus que se ha compilado para este trabajo doctoral consta de dos series completas de manuales de ILE (más de 3.600 páginas). Una

extracción pormenorizada de esta magnitud necesita tiempo y, para el presente estudio, se ha realizado en cinco meses[220]. Empleando un extractor automático se hubiera tardado aproximadamente un mes, incluyendo el análisis humano de los resultados de la extracción. Sin embargo, los resultados no hubieses sido tan exactos por las razones que se han descrito el apartado anterior. Por todo ello, y frente a una pérdida en términos cualitativos y cuantitativos de los resultados, se optó por priorizar la calidad (y la cantidad) frente a la rapidez y se optó por realizar una extracción de tipo manual cuyas fases se describen a continuación.

Los manuales que componen el corpus han sido analizados en su totalidad. En este sentido, se ha extraído fraseología en todas sus partes, sección por sección y página por página: instrucciones para la realización de los ejercicios, lecturas, parte gramatical, ejercicios al final de cada unidad y al final del manual, anexos, parte gramatical al final del libro si presente, glosarios, etc. Lo único que se decidió excluir del análisis fueron los audios[221] puesto se pretendía analizar un corpus escrito frente a uno oral.

Como ya se apuntó a lo largo de este trabajo doctoral, de acuerdo con Corpas Pastor (1996, 2001, 2013), las unidades que se han extraído del corpus reflejan el universo fraseológico desde una perspectiva amplia, incluyendo todas las Esferas (Corpas Pastor, 1996) descritas pormenorizadamente en el capítulo III del presente trabajo. Asimismo, se decidió incluir también el lenguaje especializado presente en los manuales puesto que se observó su aparición sistemática a partir de los niveles superiores de los manuales objeto de estudio.

[220] La extracción manual se realizó entre los meses de junio y octubre de 2019, si bien cabe destacar que esta tarea se compaginó con el trabajo principal del doctorando en el Centro de Traducción de los Órganos de la Unión Europea (CdT).

[221] Con excepción del audio 6 de la unidad C1 del manual *Nuovo Espresso* al tratarse de el único audio específico sobre la comprensión de unidades fraseológicas.

En cambio, no se incluyó toda aquella fraseología considerada irrelevante para la didáctica del ILE y, por ende, para el presente estudio. En este sentido, se decidió no extraer:

- UF que definen construcciones gramaticales como *preposizione articolata, nome composto, aggettivo possessivo*; y nombres para identificar los tiempos verbales como *passato prossimo, futuro anteriore*, etc. Estas locuciones se omitieron puesto que dicha fraseología *per se* no se considera productiva para el aprendizaje activo de una lengua extranjera.

- Construcciones verbales compuestas por verbos más pronombre/s (pronombres directos, indirectos o las partículas *ci* y *ne* en todas sus funciones semánticas). En este sentido, verbos como *farcela*, *pensarne*, *volerne*, etc., no se han considerado como unidades fraseológicas puesto que están compuestos por un solo término y, por tanto, no son objeto de estudio de la fraseología. En cambio, se han extraído dichos verbos si formaban parte de una construcción más amplia y cristalizada como *non ne posso più!*, *avercene fin sopra i capelli; ce la faremo!*, etc.

- Locuciones que expresan nombres propios pertenecientes al campo semántico de las fiestas italianas que no son compartidas en todo el país y que, por lo tanto, no entran en el objetivo de este trabajo, es decir, el estudio de la fraseología del italiano estándar. En este sentido, se han excluído unidades como *Festa di sant'Ambrogio* (u otro/a santo/a), la *Festa del Tulipano*, la *Festa del Redentore*, etc., pero se han incluido unidades como *Tutti i santi*, *Santo Stefano*, *Vigilia di Natale,* al ser compartidas por todos los italofonos.

- Locuciones que expresan nombres propios pertenecientes al campo semántico de la comida: tipos de pasta como *pasta alla carbonara, linguine allo scoglio*, *spaghetti cacio e pepe*, etc.; nombres de platos de carne o pescado como *pollo alla cacciatora, maiale in agrodolce*, etc.; postres como *coppa del nonno, pastiera napoletana,* etc.; y así

sucesivamente. Dicha elección se justifica por la excesiva variedad de platos propios de la tradición italiana. A pesar de su notoriedad internacional en algunos casos, si se hubiera escogido sólo los platos que se consideraban comunes en todo el país se hubiera caído en un alto grado de subjetividad. Por ello, se decidió omitirlos a priori.

- Locuciones que expresan nombres propios pertenecientes al campo semántico de los animales y, en particular, las locuciones que definen su raza como *orso bruno*, *gatto persiano, pastore tedesco*, etc.
- Locuciones que expresan nombres propios pertenecientes al campo semántico de la toponomástica como *Bassano del Grappa, Milano Marittima, Valle d'Aosta,* etc.

Tras identificar la fraseología relevante para cumplir con los objetivos del presente estudio, se ha procedido a su anotación en un archivo fraseológico en fomato electrónico y elaborado *ad hoc*. A continuación, se describe la fase de anotación y transferencia de las unidades fraseológicas.

5.2.2.3. Anotación de la fraseología en el archivo fraseológico

La anotación de las unidades se desarrolla en dos fases consecutivas y paralelas: anotación y transferencia de la fraseología extraída al archivo fraseológico[222]. Por un lado, la anotación tiene como objetivo principal el registro de los términos, así como son usado en la lengua escrita o hablada. Es, por lo tanto, una fase de la labor metodológica de primaria importancia puesto que de ella dependerá la consistencia de los datos recabados. Por otra parte, en la fase de transferencia se organiza y transfiere el repertorio fraseológico extraído en la fase anterior en el archivo fraseológico de acuerdo con un método estandarizado.

[222] Con el término archivo fraseológico se entiende un archivo en-linea (formato .xlsx) que presenta un listado de UF ordenadas a partir de criterios comunes (véase apartado 1.5.).

Con el objetivo de estandarizar la anotación de las entradas fraseológicas extraídas, se ha respetado el International Standard Organization y, en particular la ISO 704: 2000, *Terminology work. Principle and methods* que estandariza la forma canónica para anotar los términos en un expediente terminológico. En este sentido, se han anotado las unidades de la siguiente manera:

- En minúscula.
- Los sustantivos y en singular excepto si su forma cristalizada es en plural.
- Los sustantivos sin artículos.
- Los verbos en su forma infinitiva.
- En función del orden natural de aparición de componentes que conforman la unidad fraseológica.
- Con la ortografía propia del italiano estándar.
- Cada unidad se anota solo una vez y en correspondencia del nivel en el que se encontró por primera vez.

Por otra parte, la anotación de las unidades sigue el criterio uninocional[223] según el cual a cada entrada corresponde un solo concepto. De este modo, las UF polisémicas disponen de una entrada propia. Cabe destacar que, debido a las necesidades implícitas del estudio, las entradas no han sido registradas según su disposición alfabética (ISO 1087-a, 2000: 13), sino en función del nivel en el que se han encontrado y, dentro de este, respetando la aparición progresiva en el manual.

La anotación de las unidades recopiladas durante la extracción manual se realizó a través de Microsoft Excel, una de las herramientas más ampliamente utilizadas para la gestión de proyectos terminológicos. La elección de dicha herramienta frente a otras más especializadas, por ejemplo MultiTerm de RWS Group (antes SDL Trados), no es baladí. El objetivo de la presente investigación no es la creación de una ficha

[223] Véase el apartado 5.1. del presente capítulo.

fraseológica completa para cada UF en la que se incluyan atributos como el contexto, la definición, las referencias, las notas, etc., sino anotar la información necesaria para alcanzar los objetivos principales de este trabajo doctoral, y por otro lado, describir la fraseología anotada como se define a continuación.

Se han creados dos archivos fraseológicos, uno por cada serie completa de manual analizado. Cada uno es por lo tanto asociado a un método didáctico de ILE y contiene los seis niveles de referencia del MCER (A1 - C2). En la línea 1 del Excel se presenta el nivel al que el manual se refiere. Cada nivel, a su vez, se divide en doce columnas en las que se anotaron[224]:

1. Las UF según su función oracional, es decir, el tipo de unidad fraseológica en relación con la taxonomía propuesta por Corpas Pastor (1996).

2. El número de página en el que se ha encontrado cada unidad.

3. El apartado del manual del que se extrajo la fraseología. Para definir dichas secciones de los manuales se han empleado cuatro letras: C - M - A - E, a saber,

- *Consegna* (C): instrucciones para la realización de los ejercicios.
- *Manuale* (M): secciones del manual que el profesorado emplea en el aula y entre las que se encuentran lecturas, gramática, ejercicios, etc.

[224] Para ello, el presente estudio doctoral se ha inspirado en el proyecto LAICO - Lessico per Apprendere l'Italiano. En dicho corpus se han archivado los materiales didácticos del ILE para su posterior tratamiento lexicométrico. Para cada ocurrencia individual, es posible establecer, además de su categoría gramatical y el lema de referencia, también en qué sección del texto aparece (por ejemplo, si está presente en los títulos, o en los ejercicios, etc.), cuál es su ubicación geográfica dentro del texto y, por último, en qué libro de texto se encontró. para más información véase Villarini (2013) y < https://sites.google.com/site/corpuslaico#h.kpvoldm69rc7>.

- *Allegato* (A): anexos del manual como las secciones culturales, explicaciones gramaticales, consejos y/o indicaciones para la redacción y la expresión oral, etc.
- *Esercizio* (E): ejercicios al final del manual y pensados para que el alumnado practique en casa lo aprendido en el aula.

Por último, y con el objetivo de ultimar el proceso de transferencia, se han respetado los pasos de verificación estandarizados por la ISO 1087-2, T*erminology work -- Vocabulary -- Part 2: Computer applications*, referente en materia de gestión de la terminología y control de calidad[225]. Dichos pasos se realizaron con el objetivo de detectar errores a posteriori en la base de datos elaborada. A saber:

a) Double-entry check

«Data validation to determine whether the same data entry has been stored more than once»[226] (ISO 1087-2). Se ha verificado que la misma UF no se haya anotado dos o incluso más veces. Para ello, las unidades homónimas se han analizado en el contexto en el que se han encontrado con el objetivo de eliminar las que, desde una perspectiva semántica, se repetían y evitar el fenómeno de colisión homonímica[227] (Penadés, 2015). Asimismo, se registraron unidades homónimas solo si estas presentaban diferentes conceptos y/o diferentes funciones oracionales. Para ello, se ha recurrido en primer lugar a la función del formato condicional de Excel para buscar y resaltar datos duplicados[228] y, en segundo lugar, a SynchroTerm de Terminotix para identificar eventuales unidades perdidas. Al encontrar casos de duplicados, se ha anotado solo la primera unidad fraseológica en función del orden de aparición por nivel y por página.

[225] En su original «Terminology management and Quality assurance». Traducción propia.
[226] «Validación de datos para comprobar si la misma entrada ha sido almacenada más de una vez». Traducción propia.
[227] Descrito anteriormente en el apartado 2.2.2. Identificación y extracción manual de la fraseología.
[228] Para más información véase la guía en línea de Excel: <https://support.office.com/es-es/article/buscar-y-eliminar-duplicados-00e35bea-b46a-4d5d-b28e-66a552dc138d>.

b) Consistency check

«Data validation to check whether interdependent entries comply with predefined conditions»[229] (ISO 1087-2). En este sentido, se ha verificado que en su anotación se haya respetado las pautas estandarizadas por la ISO 704: 2000, *Terminology work. Principle and methods*, y definidas al principio del presente apartado.

c) Spelling check

«Data validation to determine that all words comply with predefined spelling rules»[230] (ISO 1087-2). En la trascripción manual, aunque en la automática también, son frecuentes las erratas por lo que es necesario revisar atentamente el archivo y realizar un control ortográfico (*spell-check*). Además, es preferible utilizar un único sistema ortográfico de acuerdo con los cánones impuesto o sugerido por la academia de la lengua objeto de estudio. Aun así, pueden darse casos de regionalismos que se han destinado a una entrada específica según el principio uninocional previamente descrito.

d) Completeness check

«Data validation to see whether data are present where required»[231] (ISO 1087-2). Debido a las características de la presente investigación, las únicas informaciones complementarias que se ha anotado es el nivel de referencia MCER, la página y el apartado del manual en el que se han encontrado. En este sentido, no ha sido necesario definir ninguna de las categorías gramaticales o semánticas que se emplean en terminología para especificar el concepto (definición, género, número, contexto, fuente, referencias, etc.).

[229] «Validación de datos para comprobar si las entradas interdependientes cumplen con las condiciones predefinidas». Traducción propia.
[230] «Validación de datos para determinar que todas las palabras cumplen con las reglas de ortografía predefinidas». Traducción propia.
[231] «Validación de datos para comprobar la presencia de datos donde se requieren». Traducción propia.

e) Format check

«Data validation to determine whether data conform to a specific format»[232] (ISO 1087-2). Averiguar que los datos sean transferidos en su correcta base de datos es fundamental para el desarrollo de cualquier método científico. En este caso, al anotar las UF en función de los niveles del MCER fue fundamental no errar en la inserción de las unidades fraseológicas en su correcta columna, así como anotar congruamente la página y el apartado del manual en el que se han encontrado.

f) Plausibility check

«Data validation to determine whether data conform to a specified criteria»[233] (ISO 1087-2). En la creación de una base terminológica común es necesario confirmar que las categorías gramaticales (género, número, etc.), las definiciones, fuentes y todo el aparato descriptivo del concepto sigan las reglas que se han establecido para la constitución del modelo terminológico empleado.

Tras anotar las unidades fraseológicas, se ha recurrido a una herramienta elaborada ad hoc para averiguar la frecuencia de dichas unidades frente a un corpus italiano de lengua general. A continuación, se describe dicha fase de la metodología.

5.2.3. Presentación de la fraseología extraída

A continuación se presenta la fraseología que compone los archivos fraseológicos extraída y analizada en función del método previamente descrito.

[232] «Validación de datos para determinar si los datos se ajustan a un formato específico». Traducción propia.
[233] «Validación de datos para determinar si los datos se ajustan a un criterio específico». Traducción propia.

5.2.3.1. Nuovo Espresso

En el método Nuovo Espresso se ha extraído un total 2217 unidades fraseológicas así divididas: 343 colocaciones, 1487 locuciones, 332 formulas rutinarias y 55 paremias. Como la mayoría de los manuales de ILE (véase apartado 3.5.), este también se centra principalmente en la didáctica de las locuciones frente a otras unidades.

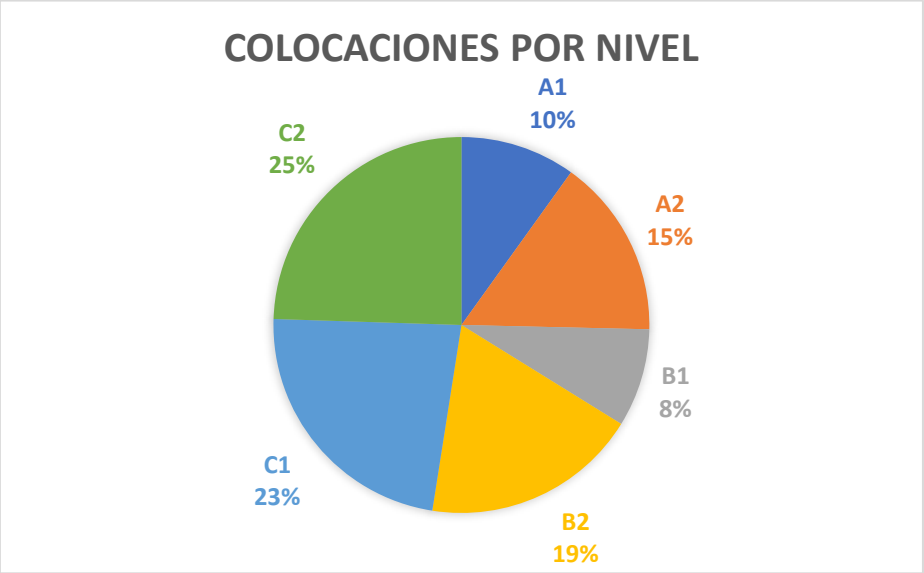
Tabla 2. Composición fraseológica de Nuovo Espresso.

	Colocaciones	**Locuciones**	**Paremias**	**Fórmulas**
A1	34	147	0	63
A2	53	191	1	91
B1	29	137	5	58
B2	64	277	10	55
C1	79	369	17	54
C2	84	366	22	11
TOT	343	1487	55	332

Fuente: elaboración propia.

Por otra parte, la distribución de las colocaciones por nivel ha sido bastante equilibrada: la mayoría de las colocaciones encontradas ha sido en los niveles superiores del MCER, C2 (25%) y C1 (23%). Mientras que en los niveles intermedios se encontraron el 19% en el B2, y el 8% en el B1. Finalmente, los niveles más básicos presentan también un significativo número de colocaciones: en el A1 el 10% mientras que en el A2 el 15%.

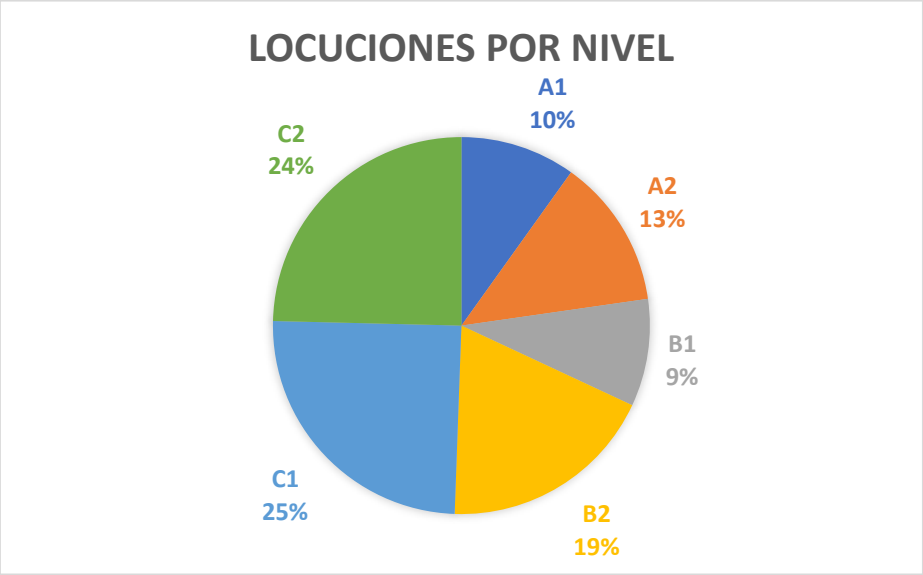
Figura 8. Distribución de las colocaciones por nivel en Nuovo Espresso.



Fuente: elaboración propia.

Asimismo, los resultados son similares con respecto a las locuciones: el mayor número de locuciones se encuentra en los niveles superiores del MCER (en el C2 el 24% y en el C1 el 25%). En cuanto a los niveles intermedios el 9% se ha encontrado en el B1 y el 19% en el B2. Finalmente, en los niveles más básicos el 10% de las locuciones se ha encontrado en el A1 y el 13% en el A2. Una vez más la distribución de las locuciones parece bastante equilibrada pro los niveles del MCER y ya a partir del nivel A1.

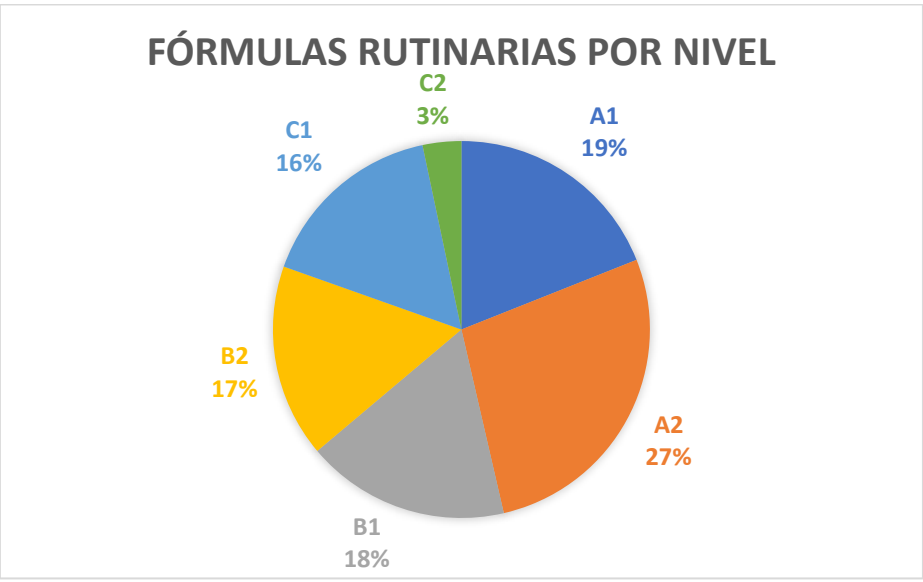
Figura 9. Distribución de las locuciones por nivel en Nuovo Espresso.



Fuente: elaboración propia.

En cuanto a las fórmulas rutinarias, se destaca que el porcentaje más alto se encuentra en el nivel A1 (27%) mientras que la menor concentración se da en el C2 (solo el 3% del total). Los demás niveles están equilibrados con porcentajes similares.

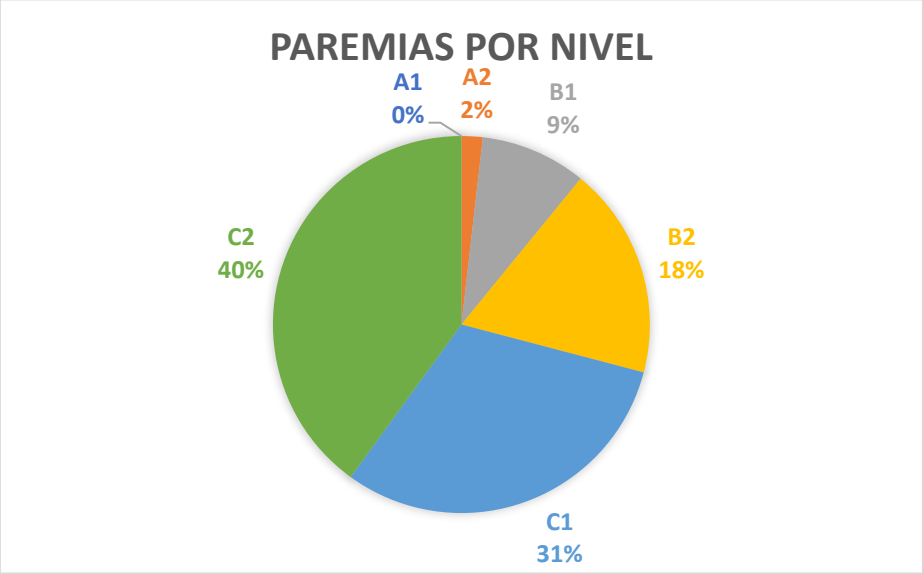
Figura 10. Distribución de las fórmulas rutinarias por nivel en Nuovo Espresso.



Fuente: elaboración propia.

Finalmente, la esfera de las paremias es duda la más desequilibrada: hasta el 70% de las paremias se encuentra distribuido en los niveles más altos del MCER: C2 (40%) y C1 (31%). El porcentaje disminuye según el nivel de modo que en el B2 aparece el 18% de las paremias y el 9% el en B1. Finalmente, cabe destacar que en el nivel A1 no se presenta ninguna paremia (el 0%).

Figura 11. Distribución de las paremias por nivel en Nuovo Espresso.



Fuente: elaboración propia.

5.2.3.2. Il Nuovo Affresco Italiano

Con respecto al segundo método, Nuovo Affresco Italiano, se han extraído un total de 2596 unidades fraseológicas. Una vez más, las locuciones representan la esfera más frecuente en este manual también con un total de 1935 locuciones presentes. Las colocaciones (380 unidades en total) y las fórmulas rutinarias (247 unidades) son las siguientes esferas más frecuentes. Finalmente, los manuales de Nuovo Affresco Italiano proponen solamente 34 paremias entre los distintos niveles los niveles.

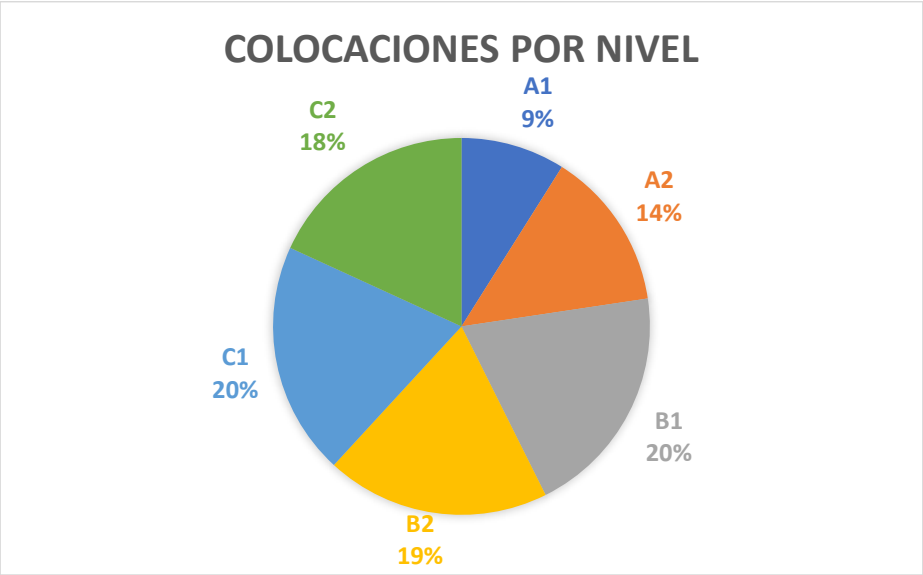
Tabla 3. Composición fraseológica de Nuovo Affresco Italiano

	Colocaciones	Locuciones	Paremias	Fórmulas
A1	34	186	0	45
A2	52	275	2	25
B1	76	446	0	29
B2	73	400	25	43
C1	76	393	1	50
C2	69	235	6	55
TOT	380	1935	34	247

Fuente: elaboración propia.

La distribución de las colocaciones por niveles en dicho método (Figura 12.) se ve principalmente repartida entre los niveles C2(18%), C1 (20%) B2 (19%) y B1 (20%). En los niveles básicos los porcentajes son menores aunque no demasiado desproporcionados (en el A1 el 9% mientras que en el A2 el 14%)

Figura 12. Distribución de colocaciones por nivel en Nuovo Affresco Italiano.



Fuente: elaboración propia.

Asimismo, con respecto a las locuciones: el mayor número de locuciones se encuentra en nivel B1 (23%), B2 (21%) y C1 (20%). Los niveles básicos presentan números similares y equilibrados: en el A1 se ha encontrado el 10% del total, mientras que en el A2 el 14%.

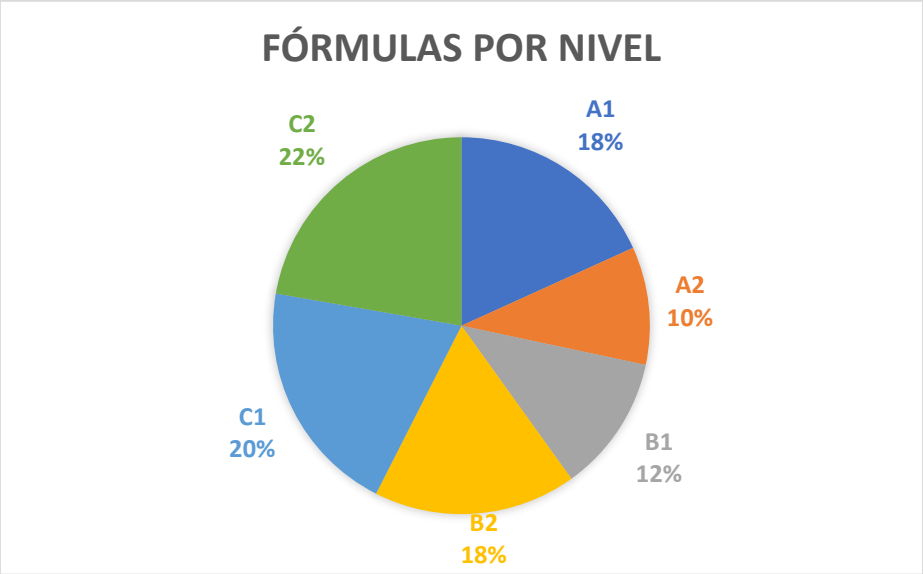
Figura 13. Distribución de locuciones por nivel en Nuovo Affresco Italiano.



Fuente: elaboración propia.

La distribución de las fórmulas rutinarias es bastante equilibrada entre los niveles del MCER: el 22% en el C2, el 20% en el C1, el 18% en el B2, 12% en el B1, 10% en el A2 y el 18% en el A1.

Figura 14. Distribución de fórmulas rutinarias por nivel en Nuovo Affresco Italiano.



Fuente: elaboración propia.

Finalmente, con respecto a las paremias, se observa la misma constante que en las otras esferas: la distribución de las esferas es bastante equilibrada por niveles (véase figura 15.) e se destaca que estas aparecen incluso en el nivel A1 (el 10% del total).

Figura 15. Distribución de paremias por nivel en Nuovo Affresco Italiano.



Fuente: elaboración propia.

5.3. Tercera fase: Análisis de la frecuencia en el *Focus corpus*

Con el objetivo de analizar y contrastar las unidades fraseológicas extraídas con un corpus general de la lengua italiana, se ha decidido realizar una prueba previa con la herramienta Sketch Engine. La aplicación online ofrece diferentes corpus de gran tamaño y, en particular, para la lengua italiana el que presenta mayor dimensión es el *Italian Web 2020* (itTenTen20[234]). La versión más reciente del corpus itTenTen consta de 12.000 millones de palabras y los textos se descargaron de la web entre noviembre y diciembre de 2019 y diciembre de 2020. itTenTen es parte de la familia TenTen corpus, un conjunto más de 30 corpus compilados a través de la web, cada uno con un tamaño superior a los 10.000 millones de palabras.

Sketch Engine permite por tanto contrastar dos corpus: el propio (llamado *Reference corpus*) con el de itTenTen20 (definido como *Focus corpus*). Sin

[234] Para más información véase <https://www.sketchengine.eu/ittenten-italian-corpus/?gclid=Cj0KCQjw7ZL6BRCmARIsAH6XFDICVF2Zm5L0V3qeQSW7ctaQriZbNEHeMKM0NNzr4Zy45PQEla4zIvUaAqjEEALw_wcB>.

embargo, y tras varios intentos fallidos puesto que la aplicación devolvía resultados completamente insatisfactorios en cuanto a fraseología, se decidió contactar con el equipo de Sketch Engine para solicitar una explicación y solución[235]. Tras un intercambio de email, el equipo de soporte confirmó que no era posible analizar la frecuencia de numerosas unidades fraseológicas de manera automática, sino que había que añadir manualmente y una por una las UF en el sistema y lanzar la busqueda contra el Focus corpus. Obviamente se trata de una solución que consume demasiado tiempo por lo que se tuvo que excluir a priori. En este sentido se observan los límites evidenciados por Hidalgo-Ternero y Corpas Pastor (2020) al afirmar que «veinte años después de la publicación seminal de Sag et al. (2002), las expresiones multiverbales (MWE) siguen siendo un dolor de cabeza para los sistemas de procesamiento de lenguaje natural».[236]

Dada la imposibilidad de recurrir a un instrumento en línea existente[237], se presenta la necesidad de elaborar una herramienta *ad hoc* que permita individuar la frecuencia de una gran cantidad de unidades fraseológicas en un corpus general de la lengua. Para dicha propuesta se recurrió a la experiencia en herramientas de procesamiento de lenguaje natural (PNL), y de extracción terminológica en particular, del grupo IATE con el que el doctorando estuvo trabajando en Centro de Traducción de la Unión Europea durante ocho meses en calidad de business analyst y coordinador de la plataforma EurTerm[238].

[235] Véase intercambio de email en Anexo 1.

[236] «Twenty years after Sag et al.'s publication seminal publication, multiword expressions (MWEs) are still a pain in the neck for natural processing systems». Traducción propia.

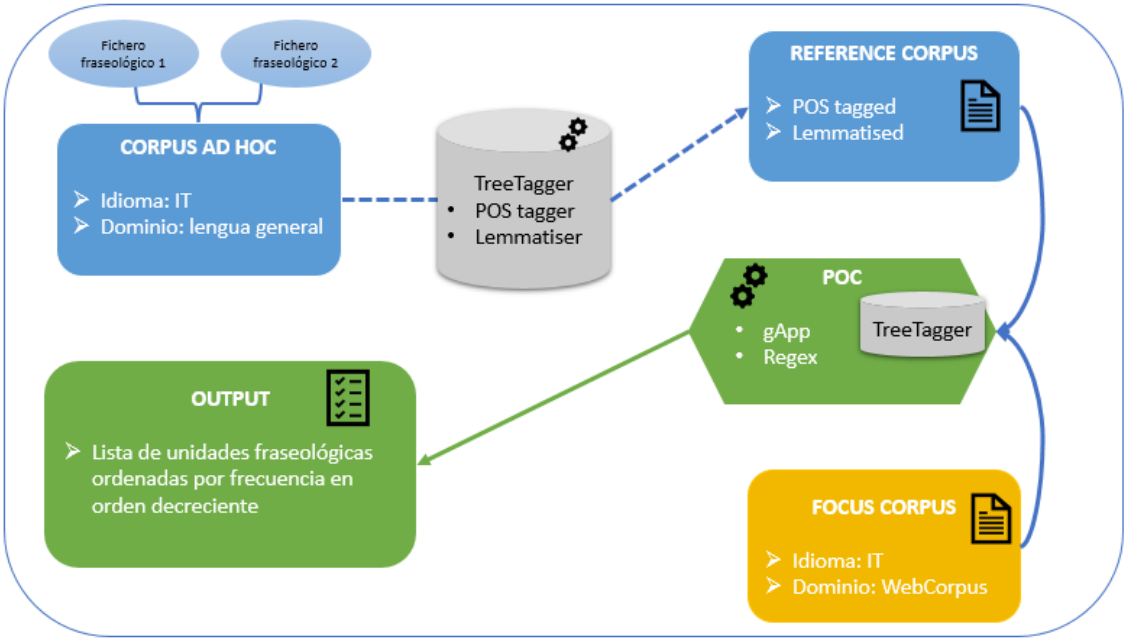
[237] Desafortunadamente los recursos del procesamiento de lenguaje natural (PLN) dedicados a la fraseología monolingüe son todavía muy escasos a demostración práctica de que «la fraseología y el PNL no han estado realmente en contacto, a pesar de ocuparse de fenómenos muy similares: unidades fraseológicas y expresiones multi-verbales» (Corpas Pastor, 2013:368).

[238] EurTerm es el portal interinstitucional de terminología creado y gestionado por el IATE Management Group. El portal fue creado con el objetivo de fomentar la cooperación en materia de terminología entre las instituciones de la UE y contiene recursos de todas las instituciones participantes, así como plataformas de colaboración para la cooperación en línea en proyectos de terminología. Para más información véase <https://termcoord.eu/intra-eu/eurterm/>.

Cabe destacar que esta tercera fase de la metodología se presenta como una propuesta en cuanto fue imposible empezar a escribir el código de la herramienta en los tiempos establecidos para defender la tesis. Tras varias reuniones con el grupo IATE se pudo poner las bases para su definición, pero, debido a las prioridades urgentes definidas por los socios interinstitucionales, al día de hoy no se ha podido todavía establecer una fecha exacta para empezar el desarrollo informático de la herramienta. Además, la situación de pandemia ocasionadas por la COVID19, el teletrabajo forzoso, y las bajas del personal han complicado ulteriormente la situación.

La Figura 16 representa una definición gráfica del proceso que permitirá identificar la frecuencia de la fraseología a través de una herramienta *ad hoc*.

Figura 16. Modelo empleado para identificar la frecuencia de las UF



Fuente: elaboración propia.

En los siguientes apartados se define cada aspecto del modelo que se empleará: se describirán los dos inputs principales, se definirá el funcionamiento de la herramienta *ad hoc*, y se describirán los riesgos y

fortalezas de dicho modelo que permitirá obtener el output deseado, una lista de unidades fraseológicas ordenadas en función de su frecuencia en un corpus general de la lengua.

5.3.1. Reference corpus

El *Reference Corpus* (RC), de acuerdo con la terminología empleada por Sketch Engine, representa el primer imput de la herramienta y está compuesto por las UF extraídas a partir de doce manuales de italiano como lengua extranjera y compiladas en formato digital (véase apartado 5.1). Para extraer de la manera más eficaz la fraseología de dicho corpus, es necesario previamente procesar las UF con el fin de: 1) obtener su anotación gramatical, y 2) lematizar el vocabulario que componen las unidades.

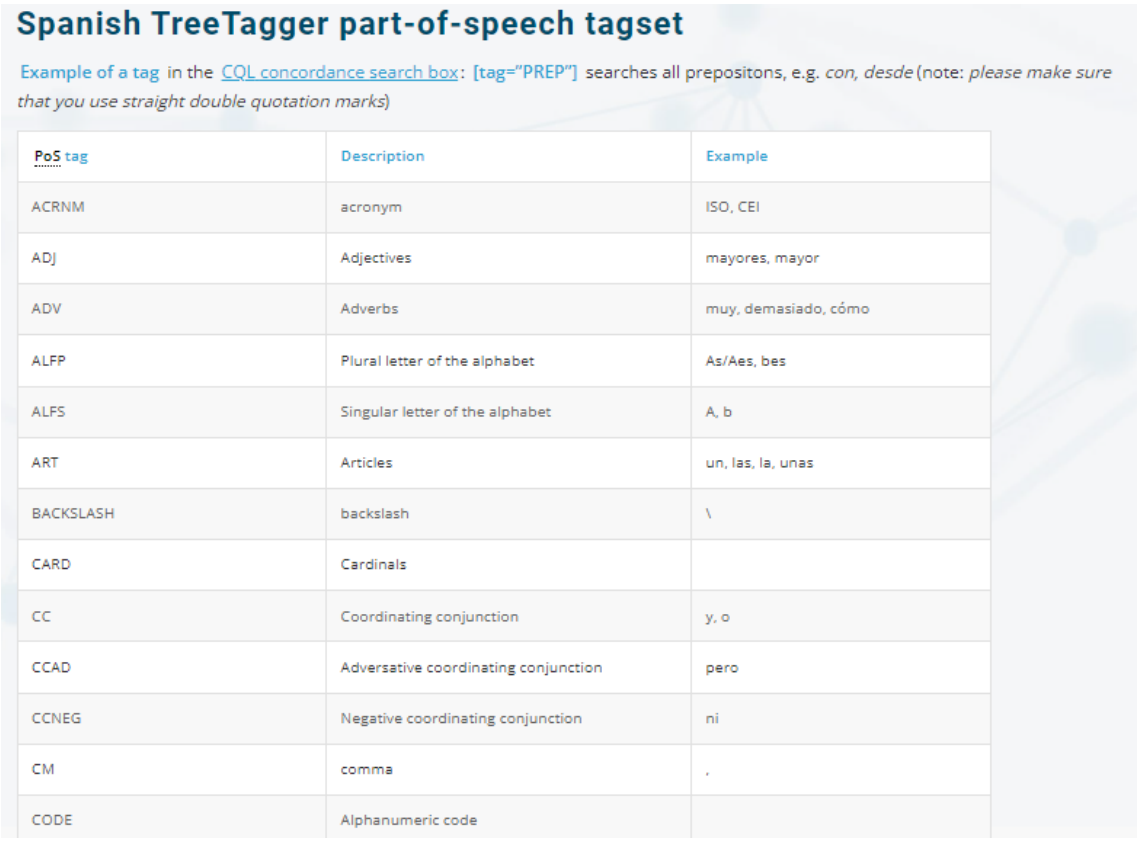
La asignación automática de descriptores morfológicos a tokens usados como input se conoce como etiquetado POS o *POS tagging* (del inglés *Part-of-Speech tagging*). En otras palabras, se trata de la asignación de categorías gramaticales (aunque no solo) a una palabra de acuerdo con su función sintáctica en el discurso. Chiche (2022) define el etiquetado POS como una importante aplicación de procesamiento del lenguaje natural que se utiliza en numerosas aplicaciones como la traducción automática, la desambiguación del sentido de las palabras, el análisis sintáctico de respuestas a preguntas, la extracción automática de términos, etc.

El mercado presenta numerosos *POS taggers*[239], herramientas informáticas que permiten anotar gramaticalmente un texto input en un idioma definido para devolver como output el mismo texto pero anotado gramaticalmente. Para su anotación automática, dichas herramientas emplean un *tag set*, es decir un conjunto de etiquetas empleadas para

[239] En función del idioma input, hay diferentes herramientas gratuitas y de pago. A modo de ejemplo véase TreTagger, SpAcy o la herramienta en línea de Natural language.Pro (<https://naturallanguage.pro/es/pos_tagging.html>).

indicar la función gramatical, y a veces también otras categorías (caso, tiempo, etc.), de cada token en un corpus dado. Por ejemplo, la etiqueta «NP» (proper name) corresponde a la categoría gramatical "nombre propio". En la figura 17 se presenta una muestra reducida del *tag set* para el español de TreeTagger[240].

Figura 17. Muestra del *tag set* de TreeTagger.



Spanish TreeTagger part-of-speech tagset

Example of a tag in the CQL concordance search box: [tag="PREP"] searches all prepositons, e.g. *con, desde* (note: *please make sure that you use straight double quotation marks*)

PoS tag	Description	Example
ACRNM	acronym	ISO, CEI
ADJ	Adjectives	mayores, mayor
ADV	Adverbs	muy, demasiado, cómo
ALFP	Plural letter of the alphabet	As/Aes, bes
ALFS	Singular letter of the alphabet	A, b
ART	Articles	un, las, la, unas
BACKSLASH	backslash	\
CARD	Cardinals	
CC	Coordinating conjunction	y, o
CCAD	Adversative coordinating conjunction	pero
CCNEG	Negative coordinating conjunction	ni
CM	comma	,
CODE	Alphanumeric code	

Fuente: Sketch Engine web page, <https://www.sketchengine.eu/spanish-treetagger-part-of-speech-tagset/>.

Empleando el *tag set* de TreeTagger, la frase «A Daniele le gusta la casa azul del final de la calle» sería anotada gramaticalmente de la siguiente manera:

[240] El *tag set* completo está disponible en: <https://www.cis.lmu.de/~schmid/tools/TreeTagger/data/spanish-tagset.txt>.

> [('A', 'PREP), ('Daniele', 'NP), ('le', 'PPC), ('gusta', 'VCLIfin), ('la' 'ART), ('casa', 'NC), ('azul', 'ADJ'), ('del', 'PREP), ('final', 'NC), ('de', 'ADP'), ('la', 'ART), ('calle', 'NP)][241]

El etiquetado de una sola palabra se basa en un sencillo algoritmo estadístico: a cada token, asigna la etiqueta que es más probable para ese token. En particular, para determinar la categoría de una palabra se usan principalmente «pistas» morfológicas, sintácticas y semánticas. En inglés se les conocen como *clue*s. El sufijo, por ejemplo, puede ser una pista morfológica significativa: si una palabra española termina en «-ción», es altamente probable que su categoría gramatical sea la de nombre, que sea un nombre común y que sea femenina. En cambio, una pista sintáctica sería considerar que un adjetivo suele aparecer mayoritariamente después de un nombre o del verbo «ser» o «estar». Finalmente, los criterios semánticos subyacen muchas de nuestras intuiciones sobre las clases de palabras, y nos permiten hacer una buena conjetura sobre la categorización de las palabras en lenguas con las que no estamos familiarizados. Por ejemplo, si sabemos que en español el sustantivo «demostracíon» en italiano se traduce con «dimostrazione», pues se puede suponer que «dimostrazione» es un sustantivo también en italiano.

Sin embargo, dicho modelo presenta ciertas limitaciones en cuanto muy limitado. Por ello, las herramientas de etiquetados se basan en lo que se conoce como *n-grams tagging*, es decir, analizan una palabra en un contexto más amplio: los tokens precedentes y los posteriores a la palabra en cuestión. La herramienta etiqueta automáticamente una palabra en función de su contexto (véase figura 18). En este sentido, en el ejemplo anterior «A Daniele le gusta la casa azul del final de la calle», la palabra «casa» será etiquetada como nombre común ('casa', 'NC), puesto que

[241] 'PREP'(Preposition), 'NP'(Proper noun), 'PPC'(Clitic personal pronoun), 'VCLIfin'(Clitic finite verb), 'ART'(Article), 'NC'(Common nouns), y 'ADJ'(Adjectives),

estadísticamente el tag anterior (t_{n-1}) atribuido a la palabra anterior (W_{n-1}) es un artículo, y el sucesivo (t_{n+1}) es un adjetivo.

Figura 18. Ejemplificación del modelo *n-grams tagging*[242]

	A	Daniele	le	gusta	la	casa	azul
Token	w_{n-5}	w_{n-4}	w_{n-3}	w_{n-2}	w_{n-1}	w_n	w_{n+1}
Tag	t_{n-5}	t_{n-4}	t_{n-3}	t_{n-2}	t_{n-1}	t_n	t_{n+1}

Fuente: elaboración propia.

En este sentido, es necesario configurar la herramienta para que asigne la etiqueta más adecuada a cada palabra, y por ello se debe definir un modelo de lenguaje específico y entrenar la máquina a través de técnicas propias de la inteligencia artificial (IA) empleada en toda herramienta de procesamiento del lenguaje natural (PNL): con cuantos más datos se entrena la máquina, mejores y más fiables serán los resultados del output.

Como se especificó anteriormente, en este trabajo no se pretende entrar en lo más profundo de las técnicas y herramientas para entrenar un *POS tagger* a través de técnicas y métodos propios del PNL, sino destacar que los *n-gram tagger* (etiquetadores de n-gramas) utilizan corpus entrenados y etiquetados para determinar qué etiqueta es la más probable en cada contexto. Dicho corpus entrenado es el que a lo largo de este trabajo, y de acuerdo con la terminología de Sketch Engine, se ha denominado *Focus Corpus*, es decir, el segundo input que nuestra herramienta utilizará para ordenar la frecuencia de las unidades fraseológicas extraídas de doce manuales de italiano como lengua extranjera y que se describirá en el siguiente apartado.

[242] Los modelos n-gramas empleados en las herramientas pueden definirse para valores mayores de n, pero cuando n es superior a 3, se ve alterado el rendimiento de los algoritmos de aprendizaje automático y su capacidad para calcular predicciones precisas (*Sparse data problem*).

Una vez que se haya realizado la fase de etiquetado morfo sintáctico o *POS tagging* a las unidades fraseológicas extraídas de doce manuales de italiano como lengua extranjera y que componen el corpus *ad hoc*, se procederá a su lematización. Lematizar es el proceso informático que consiste en asignar un lema a cada palabra de un corpus mediante una herramienta automática llamada lematizador. En otros términos, se lleva una palabra cuya forma es flexionada (genero, número, caso, conjugación, etc.) a su forma base, a la forma con la que se anotaría en un diccionario. La lematización[243] ofrece la ventaja de buscar una forma base de una palabra y obtener todas las formas derivadas, por ejemplo: la palabra «comer» en su forma infinitiva corresponde al lema del verbo; al buscar el lema «comer» en un corpus lematizado se obtendrán también todas las formas flexionadas como «comiendo», «comido», «comes», «comía», «comiera», etc.

Cabe destacar que, en parte, un proceso de lematización ya se hizo *a priori* en el acto de anotar las unidades (véase capítulo 5.1.2.3.). En este sentido, tomando como referencia el estándar internacional ISO 704: 2000, *Terminology work. Principle and methods*, se anotaron los sustantivos en su forma singular y masculina (excepto si su forma cristalizada e idiomática fuera en plural), en su forma infinitiva aquellas unidades que presentaban un verbo en su composición y, cuando necesario, se normatizó la ortografía para evitar caer en regionalismos particulares. Lematizar ulteriormente las unidades ayuda a subsanar eventuales descuidos y a reducir los riesgos en la siguiente fase de comparación con el Focus corpus.

Para su anotación semántica y su lematización se propone recurrir a una herramienta única capaz de realizar ambas fases a la vez: TreeTagger. TreeTagger es «una herramienta para la anotación automática de corpus

[243] La lematización se contrapone al *stemming*, es decir, el proceso de eliminar automáticamente de una palabra sus afijos (sufijos, prefijos, etc.) y finalmente dejar solo la raíz. El *stemming* se utiliza para detectar palabras relacionadas con la misma raíz, pero resulta ser más útil para lenguas aglutinantes mientras que para lenguas analíticas (como es el italiano) se suele preferir el método de la lematización.

con información relacionada con el etiquetado gramatical y su lematización».[244] En este sentido, procesando un corpus *ad hoc* compuesto por doce manuales de ILE con dicha herramienta es posible etiquetar automáticamente las partes del discurso y al mismo tiempo lematizar el mismo corpus. TreeTagger fue desarrollada por Helmut Schmid en el *TC Proyect* (*Textual corpora and tools for their exploration*) del Instituto de Lingüística Computacional de la Universidad de Stuttgart y se ha utilizado con éxito para etiquetar varios idiomas, como el alemán, el inglés, el francés, el italiano, el neerlandés, el español, el búlgaro, el ruso, el griego, el portugués, el chino, el swahili, el latín, el estonio y textos en francés antiguo. Cabe destacar que se eligió dicha herramienta por su versatilidad y al ser la que la plataforma IATE empleó en la creación de su *Term Extraction Module (TEM)* y su *Term Recognition Module (TRM)*. Sin embargo, cuando llegue el momento de llevar a la práctica dicha fase, se analizarán también otras posibilidades como, por ejemplo, FreeLing[245], herramienta empleada por Sketch Engine para etiquetar y lematizar su corpus itTenTen20 o spaCy, la prometedora librería gratuita y de código abierto de Pyton especializada en ofrecer numerosas herramientas NLP para más de 64 idiomas, entre los cuales figura el italiano[246].

El procesamiento de la fraseología a través de la herramienta TreeTagger dará lugar a una colección de UF gramaticalmente anotadas y lematizadas que, de acuerdo con la terminología empleada por Sketch Engine[247], en este trabajo se define como *Referece corpus*, ed decir, el corpus de referencia. Dicho corpus gramaticalmente anotado y lematizado se comparará con el *Focus corpus*, el segundo input de la herramienta y que se define a continuación.

[244] «A tool for automatic annotation of text corpora with part-of-speech and lemma information (POS tagger and lemmatizer)». Traducción propia. Para más información véase la página web del profesor Schmid <https://www.cis.lmu.de/~schmid/>.
[245] Para más información véase «https://freeling-user-manual.readthedocs.io/en/v4.1/».
[246] Para más información véase «https://spacy.io/».
[247] Para más información véase <https://www.sketchengine.eu/my_keywords/reference-corpus-2/>.

5.3.2. Focus corpus

En este trabajo se define el *Focus corpus* (FC) como el corpus del que se extraen las palabras clave y los términos. En este sentido, las unidades fraseológicas extraídas que componen el *Reference corpus* (RF) previamente descrito serán comparadas con el FC para analizar la frecuencia.

A pesar de que el *Focus corpus* que se empleará en el presente trabajo se definirá cuando se lleve a la práctica el desarrollo técnico de la herramienta, cabe abrir un paréntesis con respecto al tamaño que dicho *Focus corpus* tendrá que presentar. Para trabajar con fraseología computacional es notorio que es necesario contar con corpus de gran tamaño, muy extenso (Tschichold, 2008; Corpas Pastor, 2013; Colson, 2020; Hidalgo Ternero, 2021; etc.). Esto es debido principalmente a que las características intrínsecas de la fraseología (véase apartado 2.1.2. del presente trabajo) obligan a analizar grandes volúmenes para poder observar el comportamiento más auténtico de las unidades fraseológicas y su relación con la lengua[248]. De hecho, Corpas Pastor (2021) hace hincapié en que las UF sueles tener baja frecuencia en los corpus, por lo que es necesario trabajar con corpus de grande tamaño. Según Tschichold (2008) el tamaño ideal de un corpus compilado para estudiar la fraseología debería ser compuesto por unos cuantos centenares de millones de palabras.

Un ejemplo concreto de *Focus corpus* es el que emplea la herramienta TreeTagger. Dicho corpus fue compilado por Marco Baroni y Eros Zanchetta[249] a partir de dos corpus: el primero a partir del periódico italiano *Repubblica* (380 milliones de tokens) y, el otro, es un web corpus

[248] A modo de ejemplo, se menciona que ciertas unidades fraseológicas no son frecuentes estadísticamente hablando a pesar de ser plenamente institucionalizadas y perfectamente reconocibles por un hablante nativo. De ahí la necesidad de trabajar con corpus de gran tamaño. En este sentido, prevalece el principio del «cuanto más grande, mejor».

[249] Disponible en Creative Commons y LGPL licence en la página web <https://corpora.dipintra.it/>.

compuesto por 25 millones de token y compilado con la herramienta BootCat tools. Dicho corpus dispone de más 800 millones de tokens, más de 500 mil palabras y más de 30 mil lemas base (Schmid y Baroni, 2007), anotados gramaticalmente (POS, Part of Speech) y lematizados. Como explica el profesor Schmid en una presentación[250] realizada con Marco Baroni (2007), para lematizar y etiquetar el corpus se ha empleado la siguiente metodología:

- Lemma extraction
- Generation of inflected forms
- Manual checking
- Manual compilation of exception lists
- Different methods for the different categories we identified: verbs, adjectives, nouns, adverbs and function words

Una vez que el corpus ha sido lematizado y etiquetado semánticamente a través de un tagset[251] elaborado por el mismo Baroni[252], se ha conformado lo que, en este trabajo, y de acuerdo con la terminología empleada por Sketch Engine, se define como *Focus corpus*[253]. A pesar de que dicho corpus esté disponible en línea y sea accesible de forma libre y gratuíta, su tamaño parece no ser suficientemente adecuadas para el propósito de dicho estudio por lo que será necesario emplear otro corpus de mayores dimensiones como, por ejemplo, itTenTen disponible a través

[250] Véase <http://www.evalita.it/sites/evalita.fbk.eu/files/presentations2007/04_EVALITA_UniTn_Baroni_POS_Zanchetta.pdf>.

[251] Con tagset se hace referencia a un conjunto de etiquetas (etiquetas POS) utilizadas para indicar la parte del discurso y a veces también otras categorías gramaticales (caso, tiempo, etc.) de cada token en un corpus de texto. Para acceder al tagset empleado en dicha herramienta véase <https://home.sslmit.unibo.it/~baroni/collocazioni/itwac.tagset.txt>.

[252] El tagset de Marco Baroni es el mismo empleado por SketchEngine en las anteriores versiones de ITTenTen (2209 y 2016). Véase <https://www.sketchengine.eu/italian-treetagger-part-of-speech-tagset/>.

[253] Véase <https://www.sketchengine.eu/my_keywords/focus-corpus/>.

de la aplicación SketchEngine aunque para fines de investigación es posible solicitar su envío para ser descargado localmente[254].

Dicho corpus se ha descrito previamente (véase apartado 5.2) y consta de más de 12 mil millones de palabras recopiladas en la web en noviembre y diciembre 2019 y diciembre 2020. Una vez más, cuando se llevará a la práctica dicha fase, se realizará un segundo análisis para establecer cuál será el *Focus corpus* más adecuado para el presente estudio entre los disponibles en el mercado.

5.3.3. Propuesta para la creación de la herramienta

A continuación se describe la herramienta que se propone realizar proporcionando algunos detalles de nivel general sobre su implementación. Cabe destacar que no se pretende definir y especificar pormenorizadamente la arquitectura del sistema. El objetivo principal del presente estudio es de presentar una propuesta fraseodidáctica para identificar cuáles unidades fraseológicas enseñar en cada nivel del MCER y, en este sentido, la herramienta informática es un mero medio para alcanzar dicho objetivo final. La arquitectura final que está detrás del software de la aplicación, el *back-end* puro, la implementación del algoritmo de la máquina de vectores de soporte (del inglés *support vector machine (SVM) classifier*), así como el *fine-tuning*, será analizado y definido en colaboración con el IATE Support & Development Team y llevado a cabo como futura línea de investigación.

Python será el lenguaje de programación interpretado que se empleará para la herramienta monolingüe. Dicho lenguaje fue creado por Guido van Rossum en 1993 y en tan pocos años se ha convertido en la referencia

[254] El doctorando preguntó directamente al equipo de Sketch Engine sobre posibilidad de emplear itTenTen fuera de la aplicación y esta fue la respuesta completa: «Sketch Engine is primarily a corpus manager which enables you to search for hundreds of corpora in 90+ languages, but not to download them. Unless necessary, we do not provide full texts of the corpora, but we can do so in particular cases for research purposes (which would require you to sign a research agreement and pay a manipulation fee)».

actual para programar y trabajar con datos de grandes volúmenes (VanderPlas, 2017). El fin de dicha herramienta es ceñido y definido: encontrar la frecuencia de un número dado de unidades fraseológicas en un corpus general de la lengua. Para alcanzar un fin tan específico, una herramienta simple y linear es sin duda la elección más adecuada.

En primer lugar, utilizará como inputs el *Reference corpus* (RC) y el *Focus corpus* (FC) previamente definidos (véase apartados 5.2.1. y 5.2.2.). Dichos inputs se cargarán en la aplicación desktop directamente a través de su interfaz gráfica por medio de dos botones denominados «*Reference corpus*» y «*Focus corpus*». En este sentido, será posible testear diferentes corpus y, en una sucesiva fase de análisis, comparar los resultados empleando diferentes *Focus corpus*. El sistema de gestión de corpus que se empleará será de tipo base de datos relacionales (BDR) y, en particular, se propone emplear PostgreSQL, BDR de código abierto gestionado por el PostgreSQL Global Development Group[255].

Con el objetivo de que la herramienta se adapte a las necesidades del usuario y no al revés, se preverá cargar archivos con diferentes formatos (.docx y .xlsx, por ejemplo). Sin embargo, los corpus leíbles por máquinas están normalmente conformados por archivos de texto plano UTF8 (.txt), por lo que un módulo de la herramienta se encargará de convertir los archivos a texto plano para su posterior procesamiento. Al hacer clic en el botón «Identificar la frecuencia», la herramienta convertirá en primer lugar el formato del archivo para su análisis posterior. Los archivos originales serán conservados por la aplicación en un módulo específico con el objetivo de que el usuario pueda conservar la información original.

Una vez convertidos los inputs en texto plano UTF8, la herramienta buscará de manera asíncrona[256] en el FC cada una de las unidades fraseológicas

[255] En fase desarrollo se volverá a analizar que BDR emplear y si podrá ser más interesante utilizar un DB de tipo Corpus Query Processor (CPQ) en su lugar.

[256] Las funciones asíncronas permiten que se puedan ejecutar varios procesos al mismo tiempo evitando que se cree tráfico de datos dentro de una aplicación lo que ralentizaría todos los cálculos.

previamente lematizadas y etiquetadas que componen el RC. Para ello, el núcleo de la herramienta enviará el *Reference Corpus* al servidor TreeTagger para realizar el proceso básico de análisis del lenguaje y PLN: análisis morfológico, etiquetado POS, lematización de los términos, etc. El análisis que devolverá TreeTagger permitirá segmentar el Reference Corpus en tokens, obtener el lema de cada token y, finalmente, anotar cada token con su etiqueta morfosintáctica (*POS tagging*).

A continuación, a través de un script automático HTTP, el núcleo de la herramienta procederá con una CQL (corpus query language) *query* para analizar la frecuencia de las unidades fraseológicas del RC en el FC. Si no se encontrara una UF en todo el *Focus corpus*, se procederá a analizar la siguiente línea hasta que se encuentre una unidad fraseológica que esté en el *Focus corpus*. Cada vez que se encuentre una UF el sistema la anotará y seguirá buscando la misma unidad a lo largo de todo el FC. Cada vez que dicha unidad sea encontrada, el sistema lo anotará y seguirá buscándola hasta agotar el Focus Corpus. Agotada la búsqueda de dicha unidad, el sistema pasará a buscar la siguiente unidad y el proceso se repetirá hasta procesar todas las unidades fraseológicas que componen el *Reference Corpus*.

Finalmente, tras analizar de forma integral las UF y anotar su frecuencia en forma de tablas, la herramienta mostrará el botón «Exportar» que permitirá descargar localmente el output: un documento con formato .xlsx en el que en la columna «A» mostrará las unidades fraseológicas del RC encontradas en el FC, y en la columna «B» un valor numérico que corresponde al número de veces que dicha unidad se encontró en el FC.

5.3.3.1. Riesgos y fortalezas de la herramienta

Trabajar con macrodatos[257] (*Big Data*) presenta numerosas ventajas y fortalezas, pero, al mismo tiempo, riesgos que merece la pena destacar.

[257] Con macrodatos se hace referencia a un conjunto de datos tan voluminosos que un software tradicional de procesamiento de datos no sería capaz de gestionar. Gracias a las nuevas tecnologías. dichos

El primer riesgo de la herramienta es de tipo meramente técnico y tiene que ver con las prestaciones informáticas de la aplicación. La característica principal de ambos inputs, y del *Focus corpus* en particular, es la de presentar un conjunto de datos tan voluminosos que podrían ralentizar significativamente el análisis de los datos y la generación del output. Asimismo, el mismo hecho de cargar archivos de gran tamaño en la aplicación para su posterior procesamiento podría afectar negativamente a las prestaciones de la herramienta informática. Por todo ello, la velocidad de cálculo, el tiempo de respuesta del software en reaccionar al input (*response time*), así como el rendimiento (*performances*) de la aplicación tendrán que ser monitoreada tanto en la elaboración como en la fase de testeo de la herramienta, y habrá que gestionarla como un riesgo del proyecto con el objetivo de encontrar la mejor solución práctica para eludir o limitar dicho posible escollo.

Por otra parte, la segunda limitación que se destaca es de carácter puramente lingüístico y propia de la naturaleza semi-fija de la fraseología. Tanto en el discurso hablado como en lo escrito, las unidades fraseológicas no aparecen siempre como una secuencia contigua de palabras, sino que en ciertas unidades cabe la posibilidad de intercalar unidades sin por ello sacrificar el significado composicional o idiomático de la unidad. En la frase «mi hijo parece haber *sentado* ligeramente *la cabeza*», la unidad fraseológica *sentar la cabeza* se ve interrumpida por el adverbio *ligeramente* sin que su significado composicional se vea afectado. Sin embargo, para una herramienta de detección automática podría suponer un problema identificarla como UF, lo que afectaría negativamente a los resultados.

Numerosos estudios han analizado dicho problema y propuesto soluciones[258]. En particular, el trabajo de Hidalgo-Ternero y Corpas Pastor

volúmenes masivos de datos pueden utilizarse para resolver problemas teórico y empíricos que antes no habrían podido abordarse. Para más información véase Sagiroglu y Sinanc (2013).

[258] El trabajo de Hidalgo-Ternero y Corpas Pastor (2020) proporciona un atento y pormenorizado estado del arte en cuanto a detección de automática de unidades fraseológicas discontinuas.

ha supuesto un avance significativo en la detección automática de unidades fraseológicas discontinuas. El algoritmo ReGap, empleado en la aplicación en línea gApp, se basa en un estudio previo de patrones en el corpus y permite extraer y ordenar la fraseología detectada en un texto. En palabras de los autores, la herramienta gApp es:

> un sistema de preprocesamiento de textos basado en Python para la identificación y conversión automática de expresiones fraseológicas discontinuas a su forma continua con el fin de mejorar la traducción automática neural (NMT). (Hidalgo-Ternero y Corpas Pastor, 2020: 61)[259]

Su algoritmo, programado en Phyton, incluye también expresiones regulares (*Regex*, del inglés *Regular Expression*) creadas *ad hoc* para cada unidad fraseológicas con el objetivo de perfeccionar la detección de somatismos no consecutivos[260]. Dicha herramienta supone un gran avance en el estado de la cuestión sobre detección automática de la fraseología en cuanto por un lado es capaz de detectar las UF discontinuas en un texto y, por otro, las reordena en su forma continua.

El algoritmo se ha implementado para diferentes lenguas[261] y presenta un potencial enorme para el presente estudio en cuanto permitiría detectar unidades fraseológicas discontinuas en el *Focus corpus* y aumentar la fiabilidad del output. Cabe destacar que, como se comentó anteriormente, no sería posible implementar dicho algoritmo directamente en la herramienta, sino que sería necesario un estudio previo de cada UF que compone el *Refence Corpus*. Sin embargo, dado su potencial, al integrar dicho algoritmo en el código de la herramienta

[259] A Python-based text pre-processing system for the automatic identification and conversion of discontinuous multiword expressions (MWEs) into their continuous form in order to enhance neural machine translation (NMT). Traducción propia.
[260] El objetivo final del estudio era analizar como dicho algoritmo podía mejorar el rendimiento de los sistemas de traducción automática neuronal (NMT por sus siglas en inglés, Neural Machine Translation).
[261] A saber: español, inglés, francés, alemán, italiano y chino.

se podría limitar el riesgo de que los resultados se vean afectado por la discontinuidad de ciertas unidades fraseológicas en el *Focus corpus*.

Asimismo, desde un punto de vista semántico, otro riesgo que plantea la comparación automática de unidades fraseológicas en un corpus de gran tamaño es la colisión homonímica (Penadés Martínez, 2015), concepto ya introducido en apartados anteriores (véase 4.1.8.). Resumiendo, la colisión homonímica corresponde a la homonimia que se produce entre una locución y una combinación de palabras no fraseológica. En la frase «Daniele se levantó con el pie izquierdo», la unidad *levantarse con el pie izquierdo* puede ser una locución idiomática (levantarse de mal humor o tener mala suerte) o una combinación libre de palabras a partir del verbo *levantarse*. Si un análisis humano podría identificar su significado composicional, una herramienta informática puede llegar a tener problemas confundiendo fraseología con combinaciones libres de palabras.

Hasta el momento no se ha encontrado todavía una solución para la detección automática de casos de colisión homonímica, por lo que un análisis manual las líneas de concordancia de cada UF sería sin duda lo más adecuado. Sin embargo, dada la enorme cantidad de datos que presenta el *Focus corpus*, un análisis manual podría suponer a priori un gran esfuerzo que no devolvería un beneficio adecuado. Por ejemplo, con respecto a los somatismos, Hidalgo-Ternero y Corpas Pastor (2020) demostraron que el 50% de las unidades encontradas[262] correspondían a una combinación libre de palabras (*irrelevant result* para emplear la terminología de los autores). Si se demostrara que un porcentaje similar se aplicaría también a otros tipos de unidades diferentes a los somatismos, habría que encontrar indudablemente una solución, manual o automática, para limitar el impacto de la colisión homonímica en los resultados. En este sentido, habría que profundizar en el tema y estudiar

[262] Los autores analizan manualmente un total de 560 casos a partir de cuatro somatismos y encuentran por cada uno de ellos 70 resultados irrelevantes, es decir casos de colisión homonímica.

el riesgo pormenorizadamente en función también del tipo de unidad fraseológica.

Por otra parte, no todo son riesgos, sino que también la herramienta presenta fortalezas y ventajas significativas que se definen a continuación. La flexibilidad de la herramienta en sin duda una de sus mayores fortalezas. Como se ha visto anteriormente, la interfaz gráfica del software permitirá subir dos inputs, lo que significa que cambiando el *Focus corpus* será posible analizar como varían los resultados proporcionados y, por ende, perfeccionar el estudio.

Asimismo, se trata de una herramienta monolingüe que no se centra en ningún idioma en particular. En este sentido, se podrá emplear para realizar el mismo estudio utilizando otro idioma, lo que aportará un gran valor añadido a la fraseología y fraseodidáctica de otras lenguas. Cabe destacar que el modelo de lematización definido anteriormente es más adecuado para trabajar con lenguas analíticas (como es el italiano), mientras que para las lenguas aglutinantes se suelen preferir los procesos de *stemming*, es decir, el proceso de eliminar automáticamente de una palabra sus afijos (sufijos, prefijos, etc.) y finalmente dejar solo la raíz. Para ello habrá que analizar los resultados con idiomas aglutinantes y, si necesario, procesar los corpus con el proceso NLP de *stemming* en lugar de la lematización.

5.4. Cuarta fase: Aplicación de los IF al *Reference corpus*

La cuarta y última fase de la metodología propone la aplicación de los resultados obtenidos en la fase anterior al MCER. En el capítulo III se ha demostrado que la frecuencia de las UF es uno de los indicadores más relevantes a la hora de elegir qué unidad fraseológica enseñar en cada nivel del MCER. Por ello, el output de la tercera fase de la metodología, es decir, el listado de unidades fraseológicas italianas ordenadas en función de su frecuencia en un corpus general de la lengua, será adaptado al MCER de la siguiente manera: cuanto mayor sea la

frecuencia de una unidad fraseológica, más básico será el nivel en el que se enseñará/aprenderá.

Cabe destacar que, a pesar que la frecuencia sea el principal indicador fraseodidáctico para aprender/enseñar fraseología, tomarlo como único indicador para didactizar la fraseología de un idioma y asignarlo a un nivel del MCER presenta ciertas limitaciones (véase apartado 4.1.9.). En este sentido, se considera que la distribución final de la fraseología en función del MCER necesita considerar otros indicadores fraseodidácticos. Sin embargo, esto será parte de las futuras líneas de investigación (véase capítulo 6.2.) puesto que analizar los resultados de la frecuencia que se obtendrán a través de la herramienta *ad hoc* es propedéutico a la elección de otros indicadores.

En este estudio, se propone clasificar las UF extraídas de los manuales de ILE (segunda fase de la metodología) en función de un rango, estableciendo un nivel jerárquico según su frecuencia que permita definir el lugar que una unidad fraseológica ocupa en la lengua general. Los bloques para dividir la fraseología se proponen de la siguiente manera: a cada rango corresponde un nivel del MCER que se incrementa un 1,9 punto en función de la frecuencia al que pertenece una unidad, con excepción de los rangos 5 y 6 (C1 y C2) que se incrementan un punto cada uno. A saber:

Tabla 4. Rangos de frecuencia por nivel del MCER

Rango	**Frecuencia**	**Nivel del MCER**
Rango1	entre 10 y 8,1	A1
Rango 2	entre 8,0 y 6,1	A2
Rango 3	entre 6,0 y 4,1	B1
Rango 4	entre 4,0 y 2,1	B2

Rango 5	entre 2,0 y 1,1	C1
Rango 6	entre 1,0 y 0,1	C2

Fuente: Elaboración propia.

Los rangos 5 y 6, que corresponden a los niveles C1 y C2 del MCER, se incrementan en un solo punto por dos razones principales justificadas por la baja frecuencia de las unidades que componen dichos niveles: en primer lugar, como se ha explicado en el capítulo 3.3. sobre el «cuándo» empezar a impartir fraseología en el aula, es necesario enseñar más UF ya a partir de los niveles más básicos. En segundo lugar, al tratarse de unidades pocos frecuentes en la lengua, se hipotetiza un número superior de UF en estos dos rangos y de ahí la necesidad práctica de reducir a un punto la frecuencia.

Dicha propuesta permite asignar a cada unidad fraseológica un nivel de referencia del MCER para que tanto el profesorado como los profesionales que se dedican a la elaboración de manuales de ILE puedan tener una primera propuesta empírica a la hora de preparar sus materiales en el aula.

CAPÍTULO VI

CONCLUSIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN

6.1. Conclusiones

El presente trabajo doctoral «Indicadores fraseodidácticos: Una propuesta de sistematización de la fraseología italiana según el MCER» ha realizado una profunda revisión y análisis de las principales teorías y clasificaciones de la fraseología y de los estudios en fraseodidáctica, con especial atención a la lengua italiana y española.

La perspectiva critica que se ha empleado en dicho análisis ha permitido observar que la fraseología española y la italiana no han recibido el mismo interés académico a lo largo del tiempo. Sin embargo, se ha podido indagar en como la fraseología italiana, a pesar de que haya recibido menor interés que la española, ha mostrado una proliferación de estudios en las últimas décadas.

Asimismo, se ha podido observar que la fraseología española ha servido, y sirve todavía, de modelo para la fraseología italiana y que el mismo fenómeno de globalización e intercambio de conocimiento a nivel global favorece un clima de análisis y estudio de la fraseología italiana. En este sentido, se ha observado que hay un significativo número de académicos hispanohablantes que estudian la fraseología italiana con dedicación y cuyo análisis ha servido para favorecer el desarrollo de los estudios fraseológicos italianos en las últimas décadas.

Los italianistas han propuesto numerosas taxonomías para clasificar la fraseología italiana (véase Cap. II) aunque principalmente desde una perspectiva lexicográfica y restringida a algunos tipos de fraseologísmos. En particular, se ha constatado que la esfera de las locuciones es la que ha recibido mayor interés dentro de la fraseología italiana, hecho que encuentra demostración práctica en los manuales analizados, donde la mayoría de las unidades encontradas corresponden a la Esfera de las locuciones. Asimismo, se ha observado cierto «caos terminológico» (Mellado Blanco, 2010: 358), es decir, que la falta de una terminología aceptada por la mayoría de los italianistas ha generado una terminología inestable y variada para identificar la fraseología tanto en los estudios teóricos como en los manuales de ILE.

Por otra parte, el estudio ha profundizado en el concepto de fraseodidáctica, calco del término *phraseodidactik* acuñado por Kühn (1987), y ha adoptado la definición propuesta por González Rey (2012) que abarca una concepción lingüística, semántica y pragmática o, en otros términos, cultural de la enseñanza de la fraseología. Por todo ello, en el presente trabajo se ha empleado el término fraseodidáctica con el significado de ciencia que atañe a la enseñanza y al aprendizaje de las unidades fraseológicas, tal y como son descritas en la clasificación de Corpas Pastor (1996), desde un prisma gramatical, pragmático y semántico.

Estudiando pormenorizadamente como la fraseología italiana ha sido aplicada a la didáctica del ILE, se ha alcanzado el primer objetivo general de este trabajo, y se ha podido corroborar la primera hipótesis del presente trabajo doctoral, demostrando que la mayoría de las propuestas para aplicar los indicadores fraseodidácticos a la fraseología italiana sufren de cierto grado de subjetividad y que las colecciones de UF aplicadas a la didáctica de la fraseología no son compiladas de manera empírica, sino que se basan en gran parte en la mera experiencia subjetiva de quien los plantea.

Con el objetivo de solventar dicha ausencia, se han analizado diferentes propuestas académicas para aplicar los indicadores fraseodidácticos (IF) a la didáctica de la fraseología en el aula, lo que ha permitido alcanzar el segundo objetivo principal del presente estudio. En este sentido, se ha observado que la frecuencia es uno de los indicadores fraseodidácticos más relevantes y eficaces a la hora de enmarcar una UF en un nivel específico del MCER, lo que corrobora la segunda hipótesis del presente estudio doctoral.

El empleo de una metodología secuencial exploratoria (Creswell, 2003, 2007, 2009, 2014) ha permitido compilar un corpus compuesto por unidades fraseológicas extraídas de dos series completas de manuales de ILE. El pretest realizado en la extracción dio lugar a la comprobación de la tercera y última hipótesis, según la cual la extracción manual de la fraseología proporciona resultados más fiables que la automática, tanto desde un prisma cuantitativo como cualitativo.

Finalmente, la propuesta de un modelo para identificar la frecuencia de un corpus compuesto por unidades fraseológicas pertenecientes a diferentes esferas (Corpas Pastor, 1996) ha permitido alcanzar el tercer objetivo principal de este estudio doctoral. Dicho modelo basado en la creación de una herramienta permitirá contrastar un corpus propio compuesto de UF etiquetadas gramaticalmente y lematizadas (*Reference corpus*) con un corpus general de la lengua italiana (*Focus corpus*). El algoritmo tendrá en cuenta los riesgos relacionados con la detección automática de unidades fraseológicas discontinuas (Hidalgo-Ternero y Corpas Pastor 2020), así como el fenómeno de la colisión homonímica (Penadés Martínez, 2015).

Por todo ello se concluye que:

1. La fraseodidáctica del italiano merece recibir más atención por parte de los italianistas, puesto que la didáctica del italiano como lengua

extranjera no puede y no debe omitir la fraseología de sus programas didácticos.

2. Es necesario impartir fraseología en todos los niveles del MCER, empezando por el más básico (A1) hasta el más avanzado (C2), y establecer a priori cuáles UF enseñar en cada uno. Si bien hay una corriente teórica que considera adecuado enseñar la fraseología a partir de los niveles más altos del MCER (el mismo MCER parece adherirse), el análisis de los manuales ha demostrado que en la práctica se imparte un número significativo de unidades fraseológicas a partir del nivel A1.

3. La esfera de las locuciones es la que ha recibido mayor atención en la didáctica de la fraseología italiana como lengua extrajera. Dicha tendencia se pudo corroborar tanto en el estado del arte como en el análisis de los manuales de ILE analizados y que conforman el *Reference corpus* del presente trabajo.

4. La enseñanza/aprendizaje de las unidades fraseológicas en los distintos niveles del MCER tiene que seguir un modelo empírico que huya de subjetividades. El presente estudio es una propuesta de sistematización y no excluye su futura mejora, incluyendo más indicadores fraseodidácticos que permitan distribuir la fraseología en función del MCER.

5. La extracción automática, así como la anotación gramatical (*POS tagging*) de la fraseología a partir de un corpus dado, presenta ciertas limitaciones. Sin embargo, se han dado numerosos pasos adelante en esta dirección, y dichos avances delinean una situación prometedora y esperanzadora para el futuro de la lingüística de corpus y computacional.

6. El empleo de una herramienta informática para comparar automáticamente la fraseología presenta riesgos que hay que estudiar pormenorizadamente para que los resultados no se vean afectados. Sin

embargo, la creación de una herramienta de NLP para estudiar la fraseología en función de la frecuencia y asignarla a los distintos niveles del MCER es altamente deseable.

7. Proponiendo un método para asignar la fraseología del italiano a cada nivel del MCER, este trabajo propone subsanar una falta en la fraseodidáctica del italiano como lengua extranjera, lo que convierte el presente estudio doctoral en necesario y pertinente.

6.2. Futuras líneas de investigación

Teniendo en cuenta las conclusiones, así como las limitaciones directamente relacionadas con la naturaleza de un estudio doctoral, a continuación se plantean las futuras líneas de investigación relacionadas con el objeto de estudio.

A corto plazo, se pretende llevar a cabo el desarrollo de la herramienta informática que permita extraer la frecuencia de las UF compiladas. En colaboración con el IATE Support & Development Team se definirá pormenorizadamente la arquitectura informática que compondrá la aplicación. Aspectos como la arquitectura de sistema (*back-end*), la interfaz gráfica de usuario (*GUI*), la implementación del algoritmo de la máquina de vectores de soporte (*SVM classifier*), el *fine-tuning*, la integración de Regex en el algoritmo, etc., serán analizados y definidos en un un proyecto que se llevará a cabo colaboración con el IATE Support & Development Team.

La presente tesis doctoral es *per se* una clara apuesta por la continuidad de los estudios en fraseología y fraseodidáctica del italiano como lengua extranjera. Esta constituye el primer paso hacia un modelo más complejo de fraseodidáctica que tenga en cuenta una gama más amplia de indicadores fraseodidácticos. Por ello, a corto plazo, se afinará dicha propuesta aplicando nuevos IF al listado de UF. Con ello se pretende tener en consideración los IF que, a pesar de no ser tan significativos como la

frecuencia de aparición en la lengua general, aportarán valor añadido a la asignación de la fraseología italiana a los diferentes niveles del MCER.

En particular, se pretende empezar agregando la «Idiomaticidad y equivalencia» como indicador fraseodidáctico, ya que ha demostrado ser de los IF más relevantes para la comunidad científica. Por un lado, la idiomaticidad fraseológica, propiedad no necesaria de las UF que hace que el significado de la unidad no sea deducible por sus componentes, despierta una gran curiosidad en el alumnado y facilita la memorización de las unidades propuestas. Por otra parte, la equivalencia de ciertas unidades entre idiomas, idiomas afines en particular, facilita su fijación en el lexicón mental de un hablante, lo que se podría aprovechar en aulas homogéneas compuestas por hablantes que comparten la misma lengua materna.

Asimismo, y como objetivo a largo plazo, se pretende elaborar materiales didácticos para el italiano como lengua extranjera basados en dicha metodología. Dichos materiales supondrían un avance para el ILE ya que se basan en un corpus compilado a partir de manuales de referencia. En la elaboración de dichos manuales, se recurrirá a las valiosas aportaciones del proyecto GEPHRI (véase apartado 4.3.4.) y se propondrá colaborar con el grupo de investigación que guía este proyecto con el fin de ofrecer un valor añadido a la fraseodidáctica de la lengua italiana.

Finalmente, se pretende aplicar dicha metodología a otros idiomas romances y en particular, al español. Puesto que la herramienta informática monolingue no se basa en nigún idioma en particular, pero es capaz de extraer la frecuencia de una UF a partir de corpus, será suficiente compilar un nuevo *Reference corpus* compuesto por fraseología española y contrastarlo con un *Focus corpus* de la lengua general española. En este sentido, se podrá disponer de un listado del UF del español presentado en función del MCER, aportando un valor

añadido a la ya extensa literatura fraseológica española y dotando al profesorado de ELE de una valiosa herramienta de trabajo.

CAPÍTULO VII

REFERENCIAS BIBLIOGRÁFICAS, INDICE DE ANEXOS Y RESÚMEN EN ITALIANO

7.1. Referencias bibliográficas

Abbaticchio, Rossella (2002): "Leggere e 'sentire': Testi letterari e dimensione emotiva nell'apprendimento dell'italiano come lingua straniera". En: Guaragnella, Pasquale, *Cultura italiana, educazione linguistica*. Lecce, Pensa Multimedia, 85-104.

Alarcón Collado, Luisa y Mª Pilar, López García (1992): "¿Qué hacer con las palabras?". En: Pedro Barros García et al. (ed.), *Jornadas sobre aspectos de la enseñanza del español como lengua extranjera*. Granada, Servicio de Publicaciones de la Universidad de Granada, 33-40.

Alessandro, Arianna (2015): "Didáctica de la fraseología: una experiencia desde el marco de la investigación en la acción educativa", *Revista del Instituto de Estudios en Educación Universidad del Norte*, 22, 172-192.

Alessandro, Arianna (2011): *Investigación en la acción educativa. Las unidades fraseológicas pragmáticas en la didáctica del español y del italiano como lenguas extranjeras*. Tesis doctoral, Universidad de Murcia, Facultad de Educación, Departamento de Didáctica de la Lengua y la Literatura.

Alessandro, Arianna (2008): "La traduzione delle unità fraseologico-pragmatiche nel registro colloquiale-informale: enunciati pragmatici e Idiomatico-pragmatici in Mai sentita così bene e Historias del Kronen".

En P. R. Piras et al. (ed.), *Italianisti in Spagna, ispanisti in Italia: la traduzione*, 91-115.

Alonso Ramos, Margarita (Ed.) (2006): *Diccionarios y fraseología*. A Coruña, Servizo des Publicacións da Universidade da Coruña.

Aprile, Gianluca (2008): *Italiano per modo di dire*. Firenze, Alma Edizioni.

Austin, John l. (1996): *¿Cómo hacer cosas con palabras?*. Madrid, Paidós.

Balboni, Paolo E. (2013a): *Fare educazione linguistica: Insegnare italiano, lingue straniere e lingue classiche*. Torino, UTET Libreria.

Balboni, Paolo E. (2012): "Il ruolo delle emozioni di studente e insegnante nel processo di apprendimento e insegnamento linguistico". *Educazione linguistica,* 2, 1, 7-30.

Bagnato, Karin (2008): "La percezione dell'identità professionale". Collana di Studi Internazionali di Scienze Filosofiche e Pedagogiche. (Consutado en línea el 02/05/2020) <https://www.topologik.net/Karin_Bagnato.htm>.

Balboni, Paolo E. (2012): "Giovanni Freddi, fondatore della scuola veneziana di glottodidattica: In memoriam". *EL.LE Educazione linguistica - Language Education*, 1 (2). (Consutado en línea el 02/05/2020) <http://edizionicf.unive.it/index. php/ELLE/article/view/242>.

Balboni, Paolo E. (2011): *Verità, conoscenza ed etica nell'educazione linguistica*. Perugia, Guerra.

Balboni, Paolo E. (2010): "La glottodidattica veneziana: una 'scuola'?". En: Balboni, Paolo E.; Cinque, G. (ed.). *Seminario di Linguistica e Didattica delle Lingue: Scritti in onore degli ottant'anni di Giovanni Freddi*. Venezia, Cafoscarina.

Bally, Charles (1909): *Traité de stylistique française*. Heidelberg: Winter.

Baralo, Marta (2006): "Cómo crear redes entre palabras en el aula de ELE", *III Encuentro Práctico de Profesores de ELE Wurzburg* (Consutado en línea el 02/05/2020) < http://www.encuentropractico.com/pdfw06/baralo.pdf>.

Bardosi, Vilmos (1983): *Les locutions françaises en 150 exercices*. *Egységes jegyzet.* Budapest: Tinta Tankönyvkiadó.

Barrón-Cedeño, Alberto; Giovanni, Da San Martino, Israa, Jaradat, y Preslav Nakov (2019): "Proppy: A system to unmask propaganda in online news"*, en The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019*, Honolulu, Hawaii, 9847–9848.

Bauman, Zygmunt (2005 [2000]): *Modernidad líquida*, Fondo de Cultura Económica, Argentina.

Boas, Hans (2010): *Contrastive Studies in Construction Grammar*. Amsterdam/Philadelphia, John Benjamins. 1-20.

Borello, Enrico (2005): "Dalla competenza comunicativa alla competenza emozionale nella didattica linguistica della scuola dell'infanzia". *Rassegna italiana di linguistica applicata*. 37, 1, 107-118.

Bosque, Ignacio y Javier, Gutiérrez-Rexach (2009): *Fundamentos de sintaxis formal*. Madrid, Akal.

Cabré, María Teresa; Rosa, Estopà y Mercè, Lorente (1996): "Terminología y fraseología", T*erminología: ciencia y tecnología, V Simposio de Terminología Iberoamericana Ciudad de México*, *RITerm*. (Consutado en línea el 02/05/2021) <https://www.academia.edu/5304839/Terminolog%C3%ADa_y_Fraseolog%C3%ADa_MT_Cabre_M_LORENTE_R_ESTOP%C3%80_Actas_del_V_Simposio_de_Terminolog%C3%ADa_Iberoamericana_1996>.

Calvi, Maria Vittoria y Giovanna, Mapelli (2004): "Los marcadores 'bueno', 'pues', 'en fin' en los diccionarios de español e italiano". *Artifara*, 4, (Consutado en línea el 02/11/2020] <http://www.cisi.unito.it/artifara/rivista4/testi/marcadores.asp>.

Calvi, Maria Vittoria (1992): "Neologismi e dizionari: alcune ipotesi di confronto tra l'italiano e lo spagnolo di oggi". *Actas del Seminario Internazionale di Studi sul Lessic*o. Bologna, Clueb, 177-182

Caon, Fabio (2011): "Insegnare l'italiano attraverso la canzone". En: maraschi, Nicoletta y Fabio Caon: *Le radici e le al. Ll'italiano e il suo*

insegnamento a 150 anni dall'Unità d'Italia. Torino, UTET Libreria, 189-203.

Caon, Fabio (ed.) (2010). *Facilitare l'apprendimento dell'italiano L2 e delle lingue straniere*. Torino, UTET Libreria.

Caon, Fabio; Ongini, Vinicio (2008a): *L'intercultura nel pallone: Italiano L2 e integrazione attraverso il gioco del calcio*. Roma, Sinnos.

Caon, Fabio (2008b): *Educazione linguistica e differenziazione: Gestire eccellenza e difficoltà*. Torino, UTET Libreria.

Cardona, Mario (2012). "L'approccio cognitivo-emozionale e il visconte dimezzato". En: Caon, Fabio (ed) (2010): *Facilitare l'apprendimento dell'italiano L2 e delle lingue straniere*. Torino, UTET Libreria.

Cardona, Mario (2001)*: Il ruolo della memoria nell'apprendimento delle lingue*. Torino, UTET Libreria.

Carneado Moré, Zoila (1985): "Algunas consideraciones sobre el caudal fraseológico del español hablado en cuba". En: Carneado Moré, Zoila y María Antonia, Tristá Pérez, *Estudios de fraseología*, La Habana, Academia de ciencias de Cuba, 7-38.

Casadei, Federica (1997a): "Tra calcolabilità e caos: metafore ed espressioni idiomatiche nella semantica cognitiva". En Carapezza Marco (et al), *Linguaggio e cognizione. XXVIII Congresso Internazionale di Studi della Società di Linguistica Italiana*, 105-122.

Casadei, Federica (1997b) "Flessibilità lessico-sintattica e produttività semantica delle espressioni idiomatiche: un'indagine sull'italiano parlato". En Casadei, Federica, Giuliana, Fiorentino; Veri, Samek-Lodovici, *L' italiano che parliamo*. Santarcangelo di Romagna, FARA, 4-13.

Casadei, Federica (1996): *Metafore ed espressioni idiomatiche*, Roma, Bulzoni

Casadei, Federica (1995): "Per una definizione di "espressione idiomatica" e una tipologia dell'idiomatico in italiano", *Lingua e stile*, 2, 335-358.

Carapezza, Marco; Daniele, Gambarara y Franco Lo Piparo (1998): *Linguaggio e cognizione. Atti del XXVIII congresso della Società di linguistica italiana*, Roma, Bulzoni.

Casares, Julio (1950): *Introducción a la lexicografía moderna*, 3ª ed. Madrid, CSIC.

Castillo Carballo, María Auxiliadora y Juan Manuel García Platero (2018): "La apreciación afijal en contextos de sociolecto bajo en el habla de Sevilla", *Philologica canariensia*, 24, 77-88.

Castillo Carballo, María Auxiliadora; (2017): "La producción fraseográfica en su historia: diccionarios de locuciones y refranes". *Estudios de Lingüística del Español,* 38, 85-106.

Castillo Carballo, María Auxiliadora (2011): "Voces extranjeras: de la realidad lingüística peninsular a la americana". *Itinerarios: revista de estudios lingüisticos, literarios, históricos y antropológicos*, 13, 63-84.

Castillo Carballo, María Auxiliadora; (2002): "El universo fraseológico: algunos enfoques". *Revista de lexicografía*, 8, 25-41.

Castillo Carballo, María Auxiliadora y Juan Manuel García Platero (1993): "El léxico del dorado en Málaga". *Analecta malacitana: Revista de la Sección de Filología de la Facultad de Filosofía y Letras*, 16, 93-108.

Cavalla, Cristelle (2009): "Propositions didactiques pour l'enseignement d'éléments phraséologiques en FLE", *Revue Gramm-FLE*, 112-135.

Chiche, Alebachew y Betselot, Yitagesu (2022): "Part of speech tagging: a systematic review of deep learning and machine learning approach", en *Journal of Big Data*, 9,10, 1-25.

Chomsky, Noam (1980): *Rules and Representations*. Oxford: Blackwell.

Chomsky, Noam (1965): *Aspects of the Theory of Syntax*. Cambridge, Mass, MITPress.

Cini, Monica (2005): "La fraseologia tra teoria e pratica lessicografica", *Studi di lessicografia italiana*, 22, 283-318.

Clas, André, Polguère, Alain y Igor Mel'čuk (1995) *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve, Duculot; Montréal, AUPELF-UREF

Colson, Jean-Pierre (2020): "Computational phraseology and translation studies. From theoretical hypotheses to practical tools", En Corpas Pastor, Gloria; Jean-Pierre Colson (eds.), *Computational Phraseology*, 65-82.

Consejo de Europa (2002): *Marco Común Europeo de Referencia para las Lenguas: aprendizaje, enseñanza, evaluación*. Madrid: Instituto Cervantes-MECD, Anaya.

Conti, Andrea (2015): *I 200 Proverbi e Modi di dire italiani più usati: L'Italiano in tasca*. (Consultado en línea el 03/09/2022) <shorturl.at/kotx9>

Corpas Pastor, Gloria (2022). "You are driving me up the wall! A corpus-based study of a special class of resultative constructions", *Lexis*, 19.

Corpas Pastor, Gloria (2021): "Constructional idioms of 'insanity' in English and Spanish: A corpus-based study", *Lingua* Vol. 254, (3).

Corpas Pastor, Gloria y Jean-Pierre Colson (2020): *Computational Phraseology*. Amsterdam, John Benjamins Publishing Company.

Corpas Pastor, Gloria; HIdalgo Ternero, Carlos M. y Bautista Zambrana, María Rosario (2020): "Teaching idioms for translation purposes: a trilingual corpus-based glossary applied to Phraseodidactics (ES/EN/DE)". En Mena Martínez, Florentina (ed.) *Teaching Phraseology in the XXI Century*. New Challenges. Peter Lang.

Corpas Pastor, Gloria (2018): "Laughing one's head off in Spanish subtitles: a corpus-based study on diatopic variation and its consequences for translation". En Pedro Mogorrón (ed.) *Fraseología, Diatopía y Traducción, Colección "IVITRA Research in Linguistics and Literature*". Ámsterdam, John Benjamins. 54-106.

Corpas Pastor, Gloria (2015): "Register-Specific Collocational Constructions in English and Spanish: A Usage-Based Approach", *Journal of Social Sciences*, 11 (3), 139-151.

Corpas Pastor, Gloria (2014): "Aproximación ao estudo dos fraseoloxismos acromáticos con corpus "en negro" e os seus

equivalentes de tradución ao ruso", *Cadernos de fraseoloxía galega*, 16, 41-59.

Corpas Pastor, Gloria (2013): "Detección, descripción y contraste de las unidades fraseológicas mediante tecnologías lingüísticas". En Inés, Olza; Elvira Manero Richard (ed.): *Fraseopragmática*, Berlín, Frank and Timme, 335-373.

Corpas Pastor, Gloria (2007): "Europäismen — von Natur aus phraseologische Äquivalente? Von blauem Blut und sangre azul". En Martina Emsel y Juan Cuartero Otal (ed.): *Brücken: Übersetzen und interkulturelle Kommunikationen. Festschrift für Gerd Wotjak zum 65*. Geburtstag. Berlín, Peter Lang, 65-77.

Corpas Pastor, Gloria (2003a): *Diez años de investigación en fraseología: análisis sintáctico semánticos, contrastivos y traductológicos.* Madrid; Francfort: Iberoamericana; Vervuert.

Corpas Pastor, Gloria (2003b): "Fraseología y traducción", en Gloria Corpas Pastor (ed.): *Diez años de investigación en fraseología: análisis sintáctico-semánticos, contrastivos y traductológicos*, Madrid/ Frankfurt a.M.: Vervuert, 245-273.

Corpas Pastor, Gloria (2001a): "En torno al concepto de colocación", *Euskera*, XLVI/1: 89-108.

Corpas Pastor, Gloria (2001b): "La creatividad fraseológica: efectos semántico-pragmáticos y estrategias de traducción", *Paremia*, 10, 67-78.

Corpas Pastor, Gloria (2000): "Acerca de la (in)traducibilidad de la fraseología". En Gloria Corpas Pastor (ed.): *Las lenguas de Europa: Estudios de fraseologia, fraseografia y traduccion.* Granada, Comares, 483-522.

Corpas Pastor, Gloria (1996): *Manual de fraseología española*. Madrid, Gredos.

Corpas Pastor, Gloria; Miriam, Seghiri Domínguez; Rut, Gutiérrez Florido y Míriam Urbano Mendaña (2015): *Nuevos horizontes en los Estudios de Traducción e Interpretación (Resúmenes) / New Horizons in Translation*

and Interpreting Studies (Abstracts) / Novos horizontes dos Estudos da Tradução e Interpretação (Resumos). Ginebra, Tradulex.

Corpas Pastor, Gloria y Miriam, Seghiri Domínguez (ed.) (2016): *Corpus-based Approaches to Translation and Interpreting: From Theory to Applications*. Frankfurt, Peter Lang.

Corpas Pastor, Gloria y SEGHIRI DOMÍNGUEZ, M. (2007) "Investigar con corpus en traducción: los retos de un nuevo paradigma", *Procesamiento del Lenguaje Natural*, 39, 165-172.

Corpas Pastor, Gloria y Károly Morvay (2002): "Los estudios de Fraseología y Fraseografía en la Península Ibérica. (Breve presentación y orientación bibliográfica)". En *Anales Universitatis Scientarium Budapestinensis de Rolando Eötvös Nominatae. Sectio Lingüística*, 25, págs. 165-182.

Coseriu, Eugenio (1991): *El hombre y su lenguaje*. Madrid, Gredos.

Coseriu, Eugenio (1969): *Teoría del lenguaje y lingüística general: cinco estudios*. Madrid, Gredos.

Coseriu, Eugenio (1981): *Lecciones de lingüística general*. Madrid, Gredos.

Coseriu, Eugenio (1962): *Teoría del lenguaje y lingüística general*. Madrid, Gredos.

Cruz Diaz, Noa; Maite, Taboada; Ruslan, Mitkov (2015): "A Machine-Learning Approach to Negation and Speculation Detection for Sentiment Analysis", in *Journal of the Association for Information Science and Technology*, 67(9): 2118–2136.

Dardano, Maurizio (2009): *Costruire parole. La morfologia derivativa dell'italiano*. Bologna, il Mulino.

Dardano, Maurizio (1978): *La formazione delle parole nell'italiano di oggi. Primi materiali e proposte*. Roma, Bulzoni.

Daša Stanič (2014): "Le unità fraseologiche nei libri di testo d'italiano come lingua straniera. Analisi di libri di testo graduati secondo i livelli di competenza linguistica del Quadro Comune Europeo di Riferimento per la Conoscenza Delle Lingue (QCER)". *Linguistica*, 54.1, 167-187.

De Giovanni, Cosimo (2015): "Alcune note di riflessione sulla terminologia fraseologica in Italia". En Martinez de Carnero Calzada, Fernando y Luisa A. Messina Fajardo (ed.) *Studi di fraseologia e paremiologia*. Roma, Aracne. 159-171.

De Mauro, Tullio (2005): *La Fabbrica delle Parole*. Torino, Utet,

De Mauro, Tullio (2000): *Il dizionario della lingua italiana per il terzo millennio*. Torino, Paravia.

De Mauro, Tullio (1999): *Grande dizionario italiano dell'uso*. Torino, Utet.

De Mauro, Tullio (1967): I*ntroduzione, note e commento al Corso di linguistica generale*. Roma-Bari, Laterza.

De Mauro, Tullio; Miriam, Voghera (1996): "Scala mobile, un punto di vista sui lessemi complessi". En Benincà, Paola (et al.): *Italiano e dialetti nel tempo. Saggi di grammatica per Giulio C. Llepschy*. 99-131, Roma, Bulzoni.

De Mauro, Tullio; Federico, Mancini; Massimo, Vedovelli y Miriam, Voghera (1993): *Lessico di frequenza dell'italiano parlato*, Milano, Etas.

Di Natale, Francesco y Nadia, Zacchei (1996): *In bocca al lupo! Espressioni idiomatiche e modi di dire tipici della lingua italiana*, Perugia, Guerra Edizioni.

Dobrovol'skij, Dmitrij O. (2014): "Idiome in der Übersetzung und im zweisprachigen Wörterbuch". En Carmen Mellado Blanco (ed.): K*ontrastive Phraseologie Deutsch-Spanisch*. Tübingen, Julius Groos, 197-211.

Dobrovol'skij, Dmitrij O. (2000a): "Idioms in contrast: a functional view", en Gloria Corpas Pastor (ed.): *Las lenguas de Europa: Estudios de fraseología, fraseografía y traducción*. Granada, Comares, 367-388.

Dorbrovolskij, Dmitry O. (2000b): "La especificidad nacional y cultural en fraseología". En Pamies Bertrán, Antonio y Juan de Diós, Luque Durán (ed.): *Trabajos de lexicología y fraseología contrastivas*, 63-78.

Dorbrovolskij, Dmitry O. (1998a): "On Cultural Component in the Semantic Structure of Idioms", en P. Durčo (ed.): *Phraseology and Paremiology,* Bratislava, Akademija, 55-61.

Dorbrovolskij, Dmitry O. (1988): *Phraseologie als Objekt der Universalienlinguistik*, Leipzig, VEB Verlag Enzyklopädie.

Dorbrovolskij, Dmitry, O. y Elisabeth, Piirainen (1998): "Sobre los símbolos: aspectos cognitivos y figurados del lenguaje figurado". En Pamies Bertrán, Antonio y J. de D. Luque Durán (ed.): *Trabajos de lexicografía y fraseología contrastivas*, 29-54.

Dubois, Jean (1970): "Dictionnaire et discours didactique", *Langages*, 19, 35-47.

Durante, Vanda (ed.) (2014): *Fraseología y paremiología: enfoques y aplicaciones*, Madrid, Instituto Cervantes. (Consultado en línea el 13/09/2017) <https://cvc.cervantes.es/lengua/biblioteca_fraseologica/n5_durante/pamies.htm>.

Duro, Aldo (1986-1994): *Vocabolario della lingua italiana*. Roma, Istituto della Enciclopedia Italiana Fondata da Giovanni Treccani.

Elia, Annibale; Maurizio, Martinelli, y Emilio, D'Agostino (1981): *Lessico e strutture sintattiche*. Napoli, Liguori.

Escribano Sabio, Alicia (2019) "La traducción en los manuales de español para extranjeros en Alemania". *Sabir. International Bulletin of Applied Linguistics*, 1: 52-66

Ettinger, Stefan (2008): "Alcances e límites da fraseodidáctica. Diez preguntas clave sobre o estado actual da investigación". *Cuadernos de Fraseoloxía Galega*, *vol. 10.* 95-127.

Ettinger, Stefan (1992): "Techniques d'apprentissage des expressions idiomatiques", en Gilles, Dorion ; Franz, Joseph (eds.), *Le français aujourd'hui - une langue à comprendre - französisch heute. Mélanges offerts à Jürgen Olbert*. Frankfurt am Main, Diesterweg, 98-109.

Ettinger, Stefan (1982): "Formación de palabras y fraseología en la Lexicografía". En Stephan, Haensch Wolf; Sthefan, Ettinger y Robert, Werner (ed.): *La lexicografía. De la lingüística teórica a la lexicografía práctica*, Madrid, Gredos, 233-258.

Faloppa, Federico (2011), "Collocazioni". En Simone, Raffaele (ed.): *Enciclopedia dell'italiano*. Roma, Istituto dell'Enciclopedia Italiana – Treccani, 229-232.

Fasla, Daila (1996): "El nivel superior en la enseñanza del español como lengua extranjera: la expresión idiomática". En Teresa- G. Sibón y Magdalena Padilla (eds.): *Actas del I Simposium sobre Metodología y Didáctica del Español como Lengua Extranjera*. Sevilla, AULA2. 157-167.

Fernández Sola, Alejandro; Durán Muñoz, Isabel; y Corpas Pastor, Gloria (2015): "INTELITERM: Una herramienta innovadora para la gestión de corpus y textos para la traducción especializada". En Corpas Pastor, Gloria; Míriam, Seghiri Domínguez; Rut, Gutiérrez Florido y Míriam Urbano Mendaña: *Nuevos horizontes en los Estudios de Traducción e Interpretación (Resúmenes) / New Horizons in Translation and Interpreting Studies (Abstracts) / Novos horizontes dos Estudos da Tradução e Interpretação (Resumos)*. Ginebra, Tradulex.

Forment, Mar (2000): "Universales metafóricos en la significación de algunas unidades fraseológicas". *Revista de Lingüística Española*. 30 (2), 357-381.

García-Page Sánchez, Mario (2017): "Cuatro gatos. Sobre el valor simbólico del numeral en la fraseología". *PHRASIS, Rivista di studi fraseologici e paremiologici*. 1, 81-91.

García-Page Sanchez, Mario (2016): "Variantes léxicas correlativas", Dal Maso, Elena; Carmen Navarro. *Gutta cavat lapidem. Indagini fraseologiche e paremiologiche*. Mantova, Universitas Studiorum, 39-54.

García-Page Sáncez, Mario (2008): *Introducción a la fraseología española – Estudio de las locuciones*. Barcelona, Anthropos Editorial.

Gibbs, Raymond W. (1999): "Researching Metaphor". En L. Cameon y G. Low (ed.): *Researching and Applying Metaphor*. Cambridge, Cambridge University Press.

Gladrow, Wolfgang (1993): "Faktoren der zwischensprachlichen Äquivalenz in der Phraseologie". En Ema Krošláková y Peter Ďurčo (ed.): *Phraseology in Education, Science and Culture*. Nitra, University of Nitra, 110-121.

Goldberg, Adele E. (2013): "Constructionist Approaches". En Hoffmann, Thomas y Graeme, Trousdale (ed.): *The Oxford Handbook of Construction Grammar*. Oxford, Oxford University Press, 15-31.

Goldberg, Adele E. (2006): *Constructions at Work: The Nature of Generalization in Language*. New York, Oxford University Press.

Goldberg, Adele E. (1995): *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, University of Chicago Press.

Gómez González, Alba y Clara Ureña Tormo (2017): "Una propuesta para la enseñanza de las locuciones en la clase de ELE", *Foro de Profesores de E/LE*, 13, 313-323.

Gómez Molina, José Ramón (2000): "Las unidades fraseológicas del español: una propuesta metodológica para la enseñanza de las locuciones en la clase de E.L.E.". En María José Coperías et al. (ed.): *Aprendizaje y enseñanza de una segunda lengua, Quaderns de Filología, Estudis Lingüistics V*. Valencia, Universitat de València, 111-134.

González, Luís y Pollux Hernúñez (2004): *Las palabras del traductor. Actas del II Congreso "El español, lengua de traducción"*, Bruselas, ESLEtRA.

González Rey, Mª Isabel (2012): "De la didáctica de la fraseología a la fraseodidáctica", *Paremia*, 21, 67-84.

Gréciano, Gertrud (1984): «Pour un apprentissage des unités phraséologiques», *Nouveaux cahiers d'Allemand*, 2: 95-113.

Gross, Maurice (1981): "Les bases empiriques de la notion de prédicat semántique". *Langages*, 15, 63, 7-52.

Guia i Marín, Josep (1999): *Fraseologia i Estil: Enigmes Literaris a la València del Segle XV*. València, 'L'Estel', 3 i 4.

Häcki Buhofer, Annelies y Burger, Harald (ed.) (2007): *Phraseology in motion: Akten Der Internationalen Tagung Zur Phraseologie*, Baltmannsweiler, Schneider.

Hans Christian (2003): *A constructional approach to resultatives*, Stanford, CSLI Publications.

Harrald

Hausmann, Franz Josef (1984): «Wortschatzlernen ist Kollokationslernen. Zum Lehren und Lernen französi Wortverbindungen», *Praxis des neusprachlichen Unterricht*, vol. 31: 395-406.

Hausmann, Franz Josef (1989): "Le Dictionnaire de Collocations". En: Hausmann, Franz Josef (et al.)(eds.): *Dictionaries. Dictionnaires. Wörterbüch.* Berlin/New York, de Gruyter Mouton, 1000-1019.

Heid, Ulrich (2008): "Computational Phraseology: an overview". En Granger, Sylviane y Meunier, Fanny (eds): *Phraseology -- An interdisciplinary perspective*. Amsterdam/Philadelphia, John Benjamins, 337-360.

Heid, Ulrich (2007): "Computational Linguistics Aspects of Phraseology II". En: Burger, Harald; Dobrovol'skij, Dmitrij; Kühn, Peter y Norrick, Neal R. (eds.): *Ein internationales Handbuch der zeitgenössischen Forschung*, Vol. II. Berlin, New York, De Gruyter Mouton, 1036-1044.

Heine, Bernd (1997): *Cognitive foundation of grammar*. New York / Oxford, Oxford University Press.

Hidalgo-Ternero, Carlos Manuel y Corpas Pastor, Gloria (2020): "Bridging the "gApp": improving neural machine translation systems for multiword expression detection". En Häcki Buhofer, Annelies; Colson, Jean-Pierre; Steyer, Kathrin y Pamies Bertrán, Antonio: *Yearbook of Phraseology, vol. 11,* 61-80.

Higuera García, Marta (2009): "Aprender y enseñar léxico". *MarcoELE. Revista de Didáctica Español Lengua Extranjera*, 9, 111-126.

Higuera García, Marta (2006): *Estudio de las colocaciones léxicas y su enseñanza en español como lengua extranjera*. Madrid, Ministerio de educación y ciencia.

Higuera García, Marta (1997a): "Las unidades léxicas y la enseñanza del léxico a extranjeros". *REALE* 8, 35-49.

Higuera, Marta (1997b): "La importancia del componente idiomático en la enseñanza del léxico a extranjeros". *Frecuencia L,* 6, Madrid: Arco Libros.

Hoffmann, Thomas (2013): "Abstract Phrasal and Clausal Constructions". En Hoffmann, Thomas and Graeme, Trousdale (ed.): *The Oxford Handbook of Construction Grammar.* Oxford, Oxford University Press, 307-328.

Howart, Peter (1998): "Phraseology and Second Language Proficiency", Applied Linguistics, 19/1, 24-44.

Hymes, Dell (1972): "On communicative competence". En Pride, J. B. y Janet, Holmes (ed.), *Sociolinguistics,* Harmondsworth, Penguin, 269-285.

Imperiale, Riccardo y Schafroth, Elmar (2019): "Fraseologia italiana basata sull'uso: lessicografia digitale per apprendenti tra la Frame Semantics e la Grammatica delle Costruzioni". *Italiano LinguaDue*, 1, S. 1-28. (Consultado en línea el 13/09/2021)

Imperiale, Ricardo y Elmar, Schafroth (2016): "Fraseologia italiana per apprendenti L2 e specialisti: la pratica dell'approccio costruzionista", *RiCOGNIZIONI*. *Rivista di lingue, letterature e culture moderne*. 3, 6, 105-132. (Consultado en línea el 10/09/2017)

International Standard Organization (2000): *ISO 1087-2: Terminology work -- Vocabulary -- Part 2: Computer applications*. Geneva, International Standard Organization.

International Standard Organization (2000): *ISO 704: Terminology work. Principle and methods*, Geneva, International Standard Organization.

Iñesta Mena, Eva María y Antonio, Pamies Bertrán (2002): *Fraseología y metáfora: aspectos tipológicos y cognitivo.* Granada, Método Ediciones.

Iñesta Mena, Eva María (2010): "¿Cómo lo ves…? Competencia fraseológica en LE: acercamiento plutilingüe y multicultural" En Juana

Herrera (et all) (coords.): *Estudios sobre didácticas de las lenguas y sus literaturas: Diversidad cultural, plurilingüismo y estrategias de aprendizaje*. La Laguna, Universidad de La Laguna, 377-395.

Iraqui Jasone Cenoz (2004): "El concepto de competencia comunicativa". En Sánchez Lobato, Jesún e Isabel, Santos Gargallo: *Vademécum para la formación de profesores. Enseñar español como segunda lengua/lengua extranjera*. Madrid, SGEL, 449-465.

Irujo, Suzanne (1986): "A piece of cake: Learning and Teaching Idiom", ELT, 40-3, 236-242.

Jakobson, Roman (1975): *Lingüística y poética*, traducción de Ana Mª Gutierrez (1981). Madrid, Ediciones Cátedra.

Ježek, Elisabetta (2005): *Lessico. Classi di parole, strutture, combinazioni.* Bologna, il Mulino.

Konecny, Christine y Erica, Autelli (2013): "Assegnare fraseologismi ai livelli di competenza del QCER: un'impresa (im)possibile? Un'investigazione basata su esempi italiani e tedeschi", *XLVII Congresso Internazionale Società di Linguistica Italiana*, 56-87.

Kühn, Peter (1994): "Pragmatische Phraseologie: Konsequenzen für Phraseographie und Phraseodidaktik". En Barbara Sandig (ed.): *Europhras 92. Tendenzen der Phraseologieforschung*. Bochum, Universitätservlag Dr. N. Brockmeyer, 411-428.

Kühn, Peter (1992): "Phraseodidaktik. Entwicklungen, Probleme und Überlegungen für den Muttersprachenunterricht und den Unterricht DaF". *FluL, Fremdsprachen Lehren und Lernen*, 21, 169-189.

Kühn, Peter (1987): "Deutsch als Fremdsprache im phraseodidaktischen Dornröschenschlaf. Vorschläge für eine Neukonzeption phraseodidaktischer Hilfsmittel". *Fremdsprachen lehren und lernen*. 16, 62-79.

Larreta Zulategui, Juan Pablo (2001): *Fraseología contrastiva del alemán y el español. Teoría y práctica a partir de un corpus bilingüe de somatismos*. Frankfurt am Main: Peter Lang.

Larreta Zulategui, Juan Pablo (2003): "Aplicaciones didácticas de la fraseología contrastiva: problemas y soluciones en las fases de aprendizaje de los FrL (ejemplificados en las lenguas alemana y española)", *Es Zett*, 3. Oviedo, Universidad de Oviedo, 55-72.

Leal Riol, María jesús (2013): "Estrategias para la enseñanza y aprendizaje de la fraseología en español como lengua extranjera". *Paremia*, 22, 161-170.

Leal Riot, María Jesus (2011): *La enseñanza de la fraseología en español como lengua extranjera. Estudio comparativo dirigido a estudiantes anglófonos.* Valladolid, Universidad de Valladolid.

Leiva Rojo, Jorge Jesús (2013): "Propuesta de formulario electrónico para la revisión de unidades fraseológicas traducidas". *Educatio Siglo XXI*, 31, 2, 277-296.

Leiva Rojo, Jorge Jesús (2008): "La traducción de rasgos ortotipográficos del italiano al alemán y al español: estudio de un caso". En Luis Pegenaute, Janet Decesaris, Mercè Tricás y Elisenda Bernal (eds.): *La traducción del futuro: mediación lingüística y cultural en el siglo XXI, Vol. I. La traducción y su práctica,* Barcelona, Promociones y Publicaciones Universitarias, 227-240.

Leiva Rojo, Jorge Jesús (2005): *La recepción de la obra de Rossana Campo en España y Alemania: la traducción de los componentes oral y fraseológico en Mai sentita così bene*. Tesis doctoral, Universidad de Málaga.

Levinson, Stephen, (2000): *Significados presumibles: la teoría de la implicatura conversacional generalizada*. Madrid: Gredos.

Levinson, Stephen, (1995): "Three levels of meaning" en F. R. Palmer (ed.): *Grammar and meaning*, Cambridge, University Press, 90-115.

Lisciandro, Rosario (2018): "Las unidades fraseológicas en la enseñanza de una lengua extranjera (italiano-español)". Tesis doctoral. (Consultado en línea el 15/12/2020) <https://helvia.uco.es/xmlui/bitstream/handle/10396/16696/201800001776.pdf?sequence=1&isAllowed=y>.

López Vázquez, Lucía (2010): "La competencia fraseológica en los textos de los manuales de ELE de nivel superior". *Del texto a la lengua: la aplicación de los textos a la enseñanza-aprendizaje del español L2-LE*, Actas del XXI Congreso Internacional de ASELE. Salamanca, 531-542.

López-Jiménez, María Dolores (2013): "Multi-word lexical units in L2 textbooks". *Resla*, 26, 333-348.

Luque Durán, Juan de Dios y Manjón Pozas, Francisco José (1998): "Tipología léxica y tipología fraseológica: universales y particulares". En Luque Durán, Juan de Dios y Antonio, Pamies Bertrán (ed.): *Léxico y fraseología.* Granada, Método Ediciones, 139-154.

Luque Durán, Juan de Dios (2004a): "Aspectos universales y particulares de las lenguas del mundo". *Estudios de lingüística del español*, 21. (Consultado online el 20/01/2020) <http://elies.rediris.es/elies21/>.

Luque Durán, Juan de Dios (2004b): "Claves culturales del diseño de las lenguas: fundamentos de tipología fraseológica". *Estudios de lingüística del español*, 16. (Consultado online el 20/01/2020) <http://elies.rediris.es/elies16/Claves.html>.

Luque Durán, Juan de Dios y Antonio, Pamies Bertrán (ed.). (2005): La creatividad en el lenguaje: colocaciones idiomáticas y fraseología. Granada, Método Ediciones.

Luque Durán, Juan de Dios; Antonio, Pamies Bertrán y Francisco José, Manjón Pozas (2000): *Diccionario del insulto*. Barcelona, Ediciones Península.

Luque Durán, Juan de Dios; Antonio, Pamies Bertrán y Francisco José, Manjón Pozas (2000): *El arte del insulto. estudio lexicográfico*. Barcelona, Ediciones Península.

Manuel Seco, Olimpia Andrés, Gabino Ramo (2004): *Diccionario Fraseológico Documentado del Español Actual.* Madrid, Aguilar.

Marello, Carla (2006): "La formazione dell'insegnante di italiano L2 fra scuola e università". En Bosc, Franca; Carla Marello, Silvana Mosca: *Saperi per insegnare. Formare insegnanti di italiano per stranieri.*

Un'esperienza di collaborazione fra università e scuola. Torino, Loescher, 10-20.

Marello, Carla (1996): *Le parole dell'italiano. Lessico e dizionari*. Bologna, Zanichelli.

Marello, Carla (1993): *Lavorare sulla lingua col dizionario.* Torino, Loescher.

Marello, Carla (1994): "Lessico, lessicologia, dizionari: un panorama italiano"*, Atti del Seminario internazionale di studi sul lessico*. Forlì – San Marino, Bologna, Clueb, 161-166.

Marello, Carla (1989): *Dizionari bilingui con schede sui dizionari italiani per francese, inglese, spagnolo, tedesco*. Bologna, Zanichelli.

Martí Sanchez, Manuel (2016): "¿Qué es la competencia fraseológica? Preguntas y alguna respuesta". En Dal Maso, Elena y Carmen Navarro: *Gutta cavat lapidem. Indagini fraseologiche e paremiologiche*. Mantova, Universitas Studiorum, 55-68.

María Navarro, Paco (2018): "Estudio fraseológico del cómic Mafalda: análisis contrastivo español/francés". *Anales de Filología Francesa*, 26, 2018.

Martin, Robert (1996): "Sur les facteurs du figement lexical", en Martins-Baltar, M. (ed.), *La locution: entre langue et usages*. Fontenay-St Cloud.

Martínez Montoro, Jorge (2002): "La fraseología en J. Casares". En Susana, Pastor Cesteros y Ventura, Salazar García (ed.), *Estudios de lingüísitca*, 16. Alicante, Universidad de Alicante, 5-79.

Masini, Francesca (2007): *Parole sintagmatiche in italiano*. Roma, Università Roma Tre (Tesi doctoral).

Mel´čuk, Igor. (1993): "La phraséologie et son rôle dans l'enseignement/apprentissage d'une langue étrangère", *Revue de Linguistique Appliquée*, 92, 82-113.

Mellado Blanco, Carmen (2015a): "Parámetros específicos de equivalencia en las unidades fraseológicas del español y el alemán". *Revista de Filología de la Universidad de La Laguna*, 33, 153-174.

Mellado Blanco, Carmen (2015b): "Antiphrasis-based Comparative Constructional Idioms in Spanish". *Journal of Social Sciences*, 11 (1), 1-17.

Mellado Blanco, Carmen (Ed.)(2008): *Colocaciones y fraseología en los diccionarios*. Fráncfort, Peter Lang.

Mellado Blanco, Carmen (2010): "La "forma interna" de los fraseologismos en la teoría de Anatolij Baranov y Dmitrij Dobrovol'skij a partir de su obra Aspectos teóricos da fraseoloxía (2009)". *VERBA, Anuario Gallego de Filoxía*, 37, 345-359.

Mellado Blanco, Carmen (2004): *Fraseologismos somáticos del alemán. Un estudio léxico-semántico*. Berlin, Peter Lang.

Mellado Blanco, Carmen (1997): *Los somatismos del alemán. Semántica y estructura*, Tesis doctoral.

Mendoza Puertas, Jorge Daniel (2011): "De patitas en la calle y sin un duro. Una propuesta didáctica para la enseñanza de las unidades fraseológicas". *Revista Electrónica de Didáctica del Español como Lengua Extranjera (redELE),* 8, 21.

Meunier, Fanny; Granger, Silvian (eds.) (2008): *Phraseology in Foreign Language Learning and Teaching*. Amsterdam, John Benjamins.

Mogorrón Huerta, Pedro y Fernando, Navarro Domínguez (2015): *Fraseología, Didáctica y Traducción*. Frankfurt am Main, Peter Lang.

Mogorrón Huerta, Pedro (2015): "La variation dans les Constructions Verbales Figées de l'espagnol d'Amérique". *Linguisticae Investigationes*, 38, 2, 276-301.

Mogorrón Huerta, Pedro (2010a): "Analysis of frozenness and variance in Spanish verb constructions", *Linguisticae investigationes: Revue internationale de linguis*tique *française et de linguistique générale*, 33, 1, 86-152.

Mogorrón Huerta, Pedro (2010b): "La parasinonimia y la geosinonimia en las expresiones fijas", *Skopos: revista internacional de traducción e interpretación*, 6, 119-134.

Mogorrón Huerta, Pedro (2009): "Del origen paremiológico de algunas locuciones", *Paremia*,18, 65-76.

Mogorrón Huerta, Pedro (2002): *La expresividad en las locuciones verbales en francés y en español*. Comunidad Valenciana, Secretariado de Publicaciones de la Universidad de Alicante.

Molina García, Daniel (2006): *Fraseología bilingüe. Un enfoque lexicográfico pedagógico*. Granada, Editorial Comares.

Molina Plaza, Silavana (2004): L*a traducción de las unidades fraseológicas inglés- español: el caso de las colocaciones y las frases idiomáticas*. En González, Luís y Pollux Hernúñez (2004): *Las palabras del traductor*, *Actas del II Congreso "El español, lengua de traducción".* Bruselas, ESLEtRA, 427-443.

Montoro del Arco Esteban (2006): *Teoría fraseológica de las locuciones particulares. Las locuciones prepositivas, conjuntivas y marcadoras en español*. Frankfurt am Main, Peter Lang.

Muñoz Núñez, Mª Dolores (2011): "Reflexiones sobre el concepto de solidaridades léxicasde e. Coseriu en el marco de los recientes estudios sobre restricciones léxicasy colocaciones". En Lorenzo Hervás: *Documentos de trabajo de Lingüística teórica y general*, 109-121 (consultado en línea el 15/11/2019) <https://e-archivo.uc3m.es/bitstream/handle/10016/11540/Lorenzo_Hervas_20.pdf?sequence=5>.

Mura, G. Angela (2015): Reseña de Vera Luján, Agustín y Blanco Rodríguez, Mercedes (2014) *Cuestiones de pragmática en la enseñanza del español como 2/L*. Madrid, 107-109.

Mura, G. Angela. *(2018):* "Fraseología y fraseodidáctica digitales en E/LE". *Normas, 8, 248-258.*

Nation, P. (2001): *Learning vocabulary in another language*. Cambridge, Cambridge University Press.

Navarro, Carmen (2008a): "Didáctica de las unidades fraseológicas". En Calvi, María Victoria y Félix, San Vicente (ed.): *Didáctica del léxico y nuevas tecnologías*, Viareggio, Mauro Baroni Editore, 99-115.

Navarro, Carmen (2008b): *Aspectos de fraseología contrastiva*. Verona, Fiorini.

Navarro, Carmen, (2005): "La fraseología en los diccionarios bilingües español /italiano". En Blini, Lorenzo (et al). (ed.): *Lingüística contrastiva entre el italiano y las lenguas ibéricas / Linguistica contrastiva tra italiano e lingue iberiche*, *Actas del XXIII Congreso AISPI.* Madrid, Instituto Cervantes, 2, 228-235. (Consultado en línea el 04/07/2017) <http://www.contrastiva.it/baul_contrastivo/dati/barbero/Navarro_Fraseologia%202007.pdf>.

Navarro, Carmen (2002): "Fraseología contrastiva del español y el italiano (Análisis de un corpus bilingüe)". *Tonos digital: Revista electrónica de estudios filológicos*, 13. (Consultado en línea el 04/07/2017) <http://www.um.es/tonosdigital/znum13/secciones/estudios_U_fraseologia.htm>.

Netzlaff, Marion (2005): *La collocation adjectif-adverbe et son traitement lexicographique. Français – allemand – espagnol*. Norderstedt: Books on Demand.

Núñez Cabeza, Emilio Alejandro (2002): "Los modismos en ELE: análisis a través de los corpus digitales". En: Ana María Gimeno Sanz (ed.), *Tecnologías de la información y de las comunicaciones en la enseñanza de ELE, Actas del XII Congreso Internacional de ASELE.* Valencia, Editorial U.P.V., 159-167.

Núñez-Román, Francisco (2015): "Enseñar fraseología: consideraciones sobre la fraseodidáctica del español". *Didáctica. Lengua y Literatura*, 27, 153-166.

Olímpio de O. Silva, Mª Eugênia (2006): "Fraseología y enseñanza del español como lengua extranjera". *RedELE, 5.* (Consultado en línea el 06/10/2019) <https://www.mecd.gob.es/educacion/mc/redele/biblioteca-virtual/numerosanteriores/2006/memoriamaster/1-semestre/olimpio-d.html>.

Olímpio de Oliveira Silva, María Eugênia (2004): "El componente fraseológico en la enseñanza de E/LE", *Actas del V Congreso de Lingüística General, León,* Universidad de León, 2153-2161.

Olza Moreno, Inés (2009): *Aspectos de la semántica de las unidades fraseológicas. La fraseología somática metalingüística del español*. Tesis doctoral.

Pamies Bertrán, Antonio (2014): "El algodón no engaña: algunas observaciones sobre la motivación en fraseología". En Durante, Vanda (ed.) (2014): *Fraseología y paremiología: enfoques y aplicaciones*, Madrid, Instituto Cervantes. (Consultado en línea el 3/09/2017) <https://cvc.cervantes.es/lengua/biblioteca_fraseologica/n5_durante/pamies.htm>.

Pamies Bertrán, Antonio (2008): "Productividad fraseológica y competencia metafórica (inter)cultural". *Paremia,* 17, 41-57.

Pavel, Silvia; Diane Nolet (2001): *Manual de terminología*. Gatineau, Servicio de traducción del Gobierno de Canadá.

Penadés Martínez, Inmaculada (2017): "La enseñanza de las unidades fraseológicas". En Cestero Mancera Ana María e Inmaculada Penadés Martínez (ed.): *Manual del profesor de ELE*. Alcalá de Henares, Universidad de Alcalá, 311- 355.

Penadés Martínez, Inmaculada (2015): "La enseñanza de la fraseología vinculada a los contenidos de los manuales de ELE". En Mogorrón Huerta, Pedro y Francisco, Navarro Domínguez (ed.): *Fraseología, Traducción y Didáctica*. Frankfurt am Main, Peter Lang, 241-260.

Penadés Martínez, Inmaculada (2013): "La imagen subyacente a las locuciones como criterio de marcación". En Olza Moreno, Inés y Elvira Manero Richard (ed.): *Fraseopragmática*. Berlin, Frank & Timme, 23-47.

Penadés Martínez, Inmaculada (2012a): *Gramática y semántica de las locuciones*. Alcalá de Henares, Servicio de Publicaciones de la Universidad de Alcalá.

Penadés Martínez, Inmaculada (2012b): "Didáctica de la fraseología y de la paremiología". En Ortiz Álvarez.MariaLuiza Ortiz Alvarez (ed.): *Tendências actuáis na pesquisa descritiva e aplicada em fraseologia e paremiologia. Anais,* I, Campinas, Pontes Editores, 1-117.

Penadés Martínez, Inmaculada (2012c): "La fraseología y su objeto de estudio", *Linred*, X, 1- 17. (Consultado en línea el 13/08/2017) <http://www.linred.es/monograficos_pdf/LR_monografico10-articulo2.pdf>.

Penadés Martínez, Inmaculada (2008): *Diccionario de locuciones nominales, adjetivas y pronominales para la enseñanza del español.* Madrid, Arco/Libros.

Penadés Martínez, Inmaculada (2005): *Diccionario de locuciones adverbiales para la enseñanza del español.* Madrid, Arco/Libros.

Penadés Martínez, Inmaculada (2002): *Diccionario de locuciones verbales para la enseñanza del español*. Madrid, Arco/Libros.

Penadés Martínez, Inmaculada (2001): "Las fórmulas rutinarias: su enseñanza en el aula de E/LE". *Carabela*, 50, 83-101.

Penadés Martínez, Inmaculada (1997): "Aproximación pragmática a las unidades fraseológicas". En Escavy Zamora, Ricardo (ed.): *Homenaje a A. Roldán*. Murcia, Universidad de Murcia 411-426.

Penadés Martínez, Inmaculada (1999a): *La enseñanza de las unidades fraseológicas*. Madrid, Arco/Libros.

Penadés Martínez, Inmaculada (1999b): "Para una didáctica de las unidades fraseológicas". En Fernández González, Jesús (et al.) (ed.): *Lingüística para el siglo XXI, vol. II*, III Congreso de Lingüística General. Salamanca: Ediciones Universidad de Salamanca, 1235-1242.

Pérez Bernal, Marian (2004): "Fraseología y metáfora. Materiales para la enseñanza de la fraseología en una L2". *Las Gramáticas y los Diccionarios en la Enseñanza del Español como Segunda Lengua: Deseo y Realidad*. Actas de ASELE XV, 646-654.

Pittano, Giovanni (1992): *Frase fatta capo ha. Dizionario di modi di dire, proverbi e locuzioni.* Bologna, Zanichelli.

Plum, Alistair; Tharindu, Ranasinghe; Spencer, Jones; Constantin, Orasan; Ruslan, Mitkov (2022): *Biographical: A Semi-Supervised Relation Extraction Dataset.* (Consultado en línea el 02/05/2022) *<https://arxiv.org/pdf/2205.00806.pdf>*

Prado Aragonés, Josefina (2004): Didáctica de la lengua y la literatura para educar en el siglo XXI. Madrid, La Muralla.

Quartu, Monica y Elena, Rossi (1993): *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, Paula (2006): *Fraseología italo-española. Aspectos de lingüística aplicada y contrastiva*. Granada, Granada Lingvistica.

Rádis Baptista, Lívia Márcia Tiba (2006): "Tratándose de expresiones idiomáticas, ¡No te rompas la cabeza ni busques cinco pies al gato!". *redELE revista electrónica de didáctica / español lengua extranjera, 6*, Ministerio de Educación, Cultura y Deporte. (Consultado en línea el 02/05/2020) <http://www.educacion.gob.es/redele/revistaRedEle/2006/primera.html>.

Rakotojoelimaria, Agathe (2014): *Esbozo de un diccionario de locuciones verbales español-malgache*. Tesis doctoral, Universidad de Alcalá.

Real Academia Española (2001): *Diccionario de la lengua española* (23a ed.). Madrid, Espasa.

Recio Ariza, María Ángeles (2013): "El enfoque cognitivista en la fraseología". *Revista de Lingüística y Lenguas Aplicadas*, 8, 103-109.

Recio Ariza, María Ángeles (2011): "Gramática cognitiva y su utilidad en la enseñanza de la lengua extranjera para traductores/intérpretes". *Las vertientes de la traducción e interpretación del/al alemán*, Berlin, Frank & Timme, 163-173.

Retsker, Y. (1950): "O zakonomernykh sootvetstvijakh pri perevod'e na rodnoj jazyk". *Teorija i metodika uchebnogo perevoda*. Moskva, Izd. Akademii Pedagoguicheskij Nauk RSFSR.

Rohanian, Omid; Shiva, Taslimipoor; Samaneh, Kouchaki; Le, An Ha, y Ruslan Mitkov (2019): "Bridging the Gap: Attending to Discontinuity in Identification of Multiword Expressions". En *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1, 2692–2698.

Rothkegel, Annely (2008) "Computerlinguistische Aspekte der Phraseme I". En: En: Burger, Harald; Dobrovol'skij, Dmitrij; Kühn, Peter y Norrick, Neal R. (eds.): *Ein internationales Handbuch der zeitgenössischen Forschung*, Vol. II. Berlin, New York, De Gruyter Mouton, 1027-1035.

Ruiz Gurillo, leonor (2010): "El tratamiento de la fraseología en el Diccionario de Partículas Discursivas del Español". En Mellado, Carmen (et al) (ed.): *Nuevas perspectivas de la fraseografía del siglo XXI*. Berlin, Frank & Timme.

Ruiz Gurillo, Leonor (2008): "*¡Vaya con la mosquita muerta! Ironía para la clase de ELE*". *Revista de Didáctica MarcoELE*, 7. (Consultado en línea el 13/07/2020) <http://marcoele.com/downloads/vayaconlamosquitamuertaleonorruizgurillo.pdf>.

Ruiz Gurillo, Leonor (1997a): *Aspectos de fraseología teórica española, Anejo XXIV de Cuadernos de Filología.* Valencia, Universitat de València.

Ruiz Gurillo, Leonor (1997b): *El español coloquial y su fraseología: Los sintagmas prepositivos fraseológicos*. Tesis doctoral, Universidad de Valencia.

Ruiz Gurillo, Leonor (2002): *Ejercicios de fraseología*. Madrid, Arco/Libros.

Ruiz Gurillo, Leonor (2001a): *Las locuciones en español actual*. Madrid, Arco/Libros.

Ruiz-Gurillo, Leonor (2001b): "La fraseología como cognición: vías de análisis". *LEA: Lingüística española actual*, 23, 1, 107-132.

Ruiz Gurillo, Leonor (2000): "La fraseología", en Briz, A. y Grupo Val.Es.Co. (coords.): *¿Cómo se comenta un texto coloquial?*. Barcelona, Ariel, 169-189.

Russo, Domenico (2010): *MdD. Modi di dire. Lessico italiano delle collocazioni*, Roma, Aracne.

Sabatini, Francesco y Vittorio, Colletti (2008): *Dizionario Italiano Sabatini Coletti*, Florencia. (Consultado en línea el 02/05/2020) <http://dizionari.corriere.it/dizionario_italiano/index.shtml>.

Sag, Ivan A.; Timothy, Baldwin; Francis, Bond; Ann, Copestake y Dan Flickinger (2002): "Multiword Expressions: A Pain in the Neck for NLP". *Computational Linguistics and Intelligent Text Processing*, 1-15.

Sagiroglu, Seref y Sinanc, Duygu (2013): "Big data: A review," *International Conference on Collaboration Technologies and Systems (CTS)*, San Diego, IEEE, 42-47.

Sancho Cremades, Pelegrí (1999): *Introducció a la fraseología. Aplicació al valencià colloquial.* Valencia, Denes.

Santamaría Pérez, Isabel (2000): "Una propuesta para integrar las unidades fraseológicas en el diccionario bilingüe". En M. D. Muñoz (ed.): *Actas del IV Congreso de lingüística general*. Cádiz, Servicio de Publicaciones Universidad de Cádiz, 4, 2341-2351.

Saracho Arnáiz, Marta (2015a): "Cómo desarrollar la competencia fraseológica en la clase de ELE". *La formación y competencias del profesorado de ELE, XXVI Congreso Internacional de la ASELE.* Granada, 921-931.

Saracho Arnáiz, Marta (2015b): "Vive las expresiones, página Web". (Consultado en línea el 13/09/2020) <http://martasaracho79.wix.com/vivelasexpresiones>.

Saussure, Ferdinand (1945): *Cours de linguistique général. Por* Charles Bally e Albert Sechehaye, con la colaboracione de Albert Riedlinger, Losanna-Parigi, Payot, 1916. [Trad. it.: Ferdinand de Saussure, Corso di linguistica generale, a cura di Tullio De Mauro, Roma-Bari, Laterza, [1967].]

Schemann, Hans; Patricia, Buján; Nely, Iglesias; Juan P., Larreta; Ana, Mansilla; Carmen, Mellado (2013): *BlancoIdiomatik Deutsch-Spanisch*, Hamburg, Helmut Buske Verlag.

Seco, Manuel y Andrés, Olimpia y Gabino, Ramos (dir.) (2004): *Diccionario fraseológico documentado del español actual: locuciones y modismos españoles.* Madrid, Santillana.

Serianni, Luca (1994): "Panorama della lessicografia italiana contemporanea. En *Atti del Seminario internazionale di studi sul lessico*, Forlì-San Marino, 2/5 aprile 1992, Bologna, CLUEB, 29-43.

Serianni, Luca; Valeria, Della Valle; Giuseppe, Patota (1992): *L'italiano, la norma e l'uso della lingua*, *I suoni, i segni, le parole, il testo*. Firenze, Archimede.

Serradilla Castaño, Ana (2019): "Andando de zoca en colodra: las unidades fraseológicas con verbos de movimiento en la obra de Cervantes". *ONOMÁZEIN,* 43: 181 – 199.

Serradilla Castaño, Ana (2000): "La enseñanza de frases hechas: un método para integrar la Cultura en el aula". *Asociación para la Enseñanza del Español como Lengua Extranjera (ASELE),* Actas XI, 657-664.

Sevilla Muñoz, Julia y María, García Yelo (2008): *Las paremias en la competencia comunicativa del francés actual con vistas a la enseñanza de lenguas.* Salamanca, Ediciones Universidad de Salamanca.

Sevilla Muñoz, Julia y María Teresa, Zurdo Ruiz-Ayúcar [dir.] (2009): *Refranero multilingüe*. Madrid. Instituto Cervantes (Centro Virtual Cervantes). (Consultado en línea el 10/11/2019) <http://cvc.cervantes.es/lengua/refranero/>.

Sevilla Muñoz, Julia (2015): "Las TIC en la clase de FLE para la enseñanza de refranes y locuciones". *Fraseología, Didáctica y Traducción*. Frankfurt am Main, Peter Lang, 261-272.

Sevilla Muñoz, Julia y María Sardelli (2011): "Los falsos amigos fraseológicos", *El Trujumán.* Consultado en línea el 10/11/2019)

<https://cvc.cervantes.es/trujaman/anteriores/septiembre_11/26092011.htm>.

Sevilla Muñoz, Julia y González Rodríguez, Antonio (1995): "La traducción y la didáctica de las expresiones idiomáticas (francés-español)". *Équivalences*, 24/2, 25/1-2, 171-182.

Sevilla Muñoz (2012): "Utilización de recursos en línea en la enseñanza/aprendizaje de traducción de unidades fraseológicas". En González Rey, M.ª Isabel (ed.): *Unidades fraseológicas y TIC*. Madrid, Instituto Cervantes, 283-298.

Sevilla Muñoz, Julia y Marina, García Yelo (2008): "Las paremias en la competencia comunicativa del francés actual con vistas a la enseñanza de lenguas". *Texto, género y discurso en el ámbito francófono*, XVII Congreso de la Asociación de Francesistas de la Universidad Española (APFUE), 169-178.

Sevilla Muñoz, Julia (1997a): "Fraseología y traducción". *Homenaje al Prof J. Cantera*. Madrid, Publicaciones Universidad Complutense, 431-440.

Sevilla Muñoz, Julia (1997b): "La fraseología y la paremiología en los últimos decenios". *Linred: Lingüística en la Red*, 1-29.

Sevilla Muñoz, Julia (1993): "Las paremias españolas: clasificación, definición y correspondencia francesa". *Paremia*, 2, 15-20.

Schafroth, Elmar; Riccardo, Imperiale; Francesca, Martulli; Matthias Bürgel (2022): *GEPHRI (Gebrauchsbasierte Phraseologie des Italienischen)*. Romanistik IV, Universität Düsseldorf. (Consultado en línea el 10/05/2022) <https://gephri.phil.hhu.de/>.

Simone, Raffaele; Gioia, Insacco (2017): "Cicli lessicali e polisemia nelle nominalizzazioni. Analisi di un corpus italiano". *Studi e saggi linguistici LV* (1), 9-46.

Simone, Raffaele (2016): "Il dizionario del futuro". En Marazzini, Claudio; Ludovica Maconi (ed.): *L'italiano elettronico. Vocabolari, corpora, archivi testuali e sonori*. Firenze, Accademia della Crusca, 17-32.

Simone, Raffaele (2008): "Verbi sintagmatici come categoria e come costruzione". En Cini, Monica: *I verbi sintagmatici in italiano e nelle varietà dialettali. Stato dell'arte e prospettive di ricerca*, *Atti delle giornate di studio Torino*. Frankfurt, Peter Lang, 13-30.

Simone, Raffaele (2007), "Constructions and Categories in Verbal and Signed languages". En Pietrandrea, Paola et all. (ed.): *Verbal and Signed Languages. Comparing Structures, Constructs and Methodologies*. Mouton & De Gruyter: Berlin & New York, 197-248.

Simone, Raffaele (1997): *Fondamenti di linguistica*. Roma-Bari, Laterza.

Simone, Raffaele y Francesca, Masini (ed.) (2014): *Word Classes. Nature, typology and representations*, Amsterdam-Philadelphia, John Benjamins.

Sinclair, John y Antoinette, Renouf (1985): "A lexical learning syllabus for language". *Vocabulary and Language Teaching*. London-New York, Longman, 140–160.

Sinclair, John (1991): *Corpus, Concordance, Collocation*. Oxford, Oxford University Press.

Solodukho, E. M. (1977): Voprosy sopostavitel'nogo izuchenija zaimstvovannoj fraseologuii. Kazan', KGU.

Szyndler, Agnieszka (2015): "La fraseología en el aula de E/LE: ¿un reto difícil de alcanzar? Una aproximación a la fraseodidáctica". *Didáctica. Lengua y Literatura*, 27, 197-216.

Szyndler, Agnieszka (2014): "Sobre la pseudoequivalencia fraseológica desde una perspectiva cognitivista". *Anuario de Estudios Filólogicos*, XXXVII, 251-267.

Termignoni, Susana (2015): *Bases teórico-metodológicas para um hiperdicionário semibilíngue de expressões idiomáticas italiano-português em meio a um ambiente virtual de aprendizagem*. Tesis doctoral, Universidade Federal do Rio Grande do Sul. Instituto de Letras.

Thun, Harald (1978): *Probleme der Phraseologie. Untersuchungen zur wiederholten Rede mit Beispielen aus den Französischen, Italienischen, Spanischen und Romänischen*. Tübingen, Max Niemeyer.

Timofeeva, Larissa (2013), "La fraseología en la clase de lengua extranjera: ¿misión imposible?". *Onomázein. Revista semestral de lingüística, filología y traducción*, 28, 320-336.

Timofeeva, Larissa (2012): "Sobre la traducción fraseológica". *Estudios de Lingüística (ELUA)*, 26, 405-432.

Trier, Jost (1931): *Der deutsche Wortschatz im Sinnbezirk des Verstandes*, Thesis doctoral. Universidad de Bonn.

Tristá Pérez, Antonia María (1988): *Fraseología y contexto*. La Habana, Editorial de Ciencias Sociales.

Trivić, Aneta (2011): "Equivalencias traductológicas de ciertas locuciones del español y del serbio contemporáneo". Journal of Linguistic Intercultural Education, 4, 219-238.

Tschichold, Cornelia (2008): "A Computational Lexicography Approach to Phraseologisms". En: Granger, Sylviane y Meunier, Fanny: *Phraseology: An interdisciplinary perspective*. Amsterdam-Filadelfia, John Benjamins, 361-392.

Ureña Tormo, Clara (2017): "Una nueva propuesta didáctica para la enseñanza-aprendizaje de las locuciones y los refranes en ELE". En Ainciburu María Cecilia: *Experiencias plurilingües en la enseñanza de lenguas extranjeras*. Barcelona, Octaedro.

Ureña Tormo, Clara (2014): *La enseñanza de las locuciones en ELE*. Trabajo fin de máster. Alcalá de Henares, Universidad de Alcalá.

Vachek, John (1967): "The Non-Static Aspect of the Synchronically Studied Phonological System". *Phonologic der Gegenwart, Wiener Slavistisches Jahrbuch, Erganzungsband VI*. Graz-Viena-Koln, Bohlaus Nachf, 79, 85.

Valdés, Juan (1535): "Diálogo de la lengua". *Biblioteca virtual del Instituto Cervances* (Consultado en línea el 12/12/2020)

<http://www.cervantesvirtual.com/obra-visor/dialogo-de-la-lengua--0/html/fede437e-82b1-11df-acc7-002185ce6064_2.html>.

VanderPlas, Jake (2017): Py*thon Data Science Handbook*. Sebastopol, O'Reilly Media, Inc.

Vázquez, Mercè y Oliver, Antoni (2018). "Improving term candidates selection using terminological tokens". *Terminology*, 24,122-147.

Vietri, Simonetta (2014): The Lexicon-Grammar of Italian Idioms". *Workshop on Lexical and Grammatical Resources for Language Processing*. Dublin, Coling, 24,137-146.

Villarini Andrea (2021): Lo sviluppo della competenza lessicale: incursioni tra le ipotesi teoriche e le applicazioni glottodidattiche. En (eds): Jafrancesco E., La Grassa M., *Competenza lessicale e apprendimento dell'italiano L2*, Firenze, Firenze University Press, 51-63.

Villarini Andrea (2017): "Le tipologie didattiche sul lessico pensate per apprendenti di italiano L2 adulti", *Studi italiani di linguistica teorica e applicata*, XLVI, 157-178.

Villarini Andrea (2013): "Lo sviluppo della comptenza lessicale in italiano L2 nei manuali e nel parlato del docente", *Studi di linguistica teorica e applicata*, XLII, 3, 599-619.

Voghera, Miriam (2004): "Le poliremathiche". En Grossmann, María y Franz Rainer (ed.): *La formazione delle parole in italiano*. Tübingen, Niemeyer, 56-69.

Vietri, Simonetta (1990): "La sintassi delle frasi idiomatiche". S*tudi italiani di linguistica teorica e applicata*, XIX, 133-146.

Voghera, Miriam (1994), "Lessemi complessi: percorsi di lessicalizzazione a confronto". *Lingua e Stile*, 29, 2, 185-214.

Vietri, Simonetta (1985): *Lessico e sintassi delle espressioni idiomatiche. Una tipologia tassonomica dell'italiano*. Napoli, Liguori Editore.

Vyshnya, Natalia (2006): "Adquisición de las competencias lingüísticas en las clases del idioma español (L2) a través de las paremias clásicas de El Quijote". *Didáctica (Lengua y Literatura)*, 18, 315-329.

Wahl, Alexander y Gries, Stefan (2018): "Multi-word Expressions: A Novel Computational Approach to Their Bottom-Up Statistical Extraction: Advances and Applications". *Lexical Collocation Analysi,* 85-109.

Wierzbicka, Anna (1997a): *Semantics. Primes and Universals*. Oxford -New York, Oxford University Press.

Wierzbicka, Anna (1997b): *Understanding Cultures through their Keywords*. Oxford, Oxford University Press.

Wierzbicka, Anna (1992): *Semantics, Culture, and Cognition*. Oxford, Oxford University Press, 1992.

Wierzbicka, Anna (1972): *Semantic Primitives*. Frankfurt, Athenäum.

Woodward, T. (2001): *Planificación de clases y cursos*. Madrid, Cambridge University Press.

Wotjak, Gerd (1983): "En torno a la traducción de unidades fraseológicas (con ejemplos tomados del español y el alemán)". *Linguistische Arbeitsberichte*, 40, 56-80.

Xatara, C.M. (2001): "O ensino do léxico: as expressões idiomáticas". *Trabalhos em Lingüística Aplicada*. Campinas, IEL-UNICAMP, 37, 49-69.

Zamora Muñoz, Pablo; Arianna, Alessandro; Eleonora, Ioppoli y Federica, Simone (2006): *Hai voluto la bicicletta... Esercizi su fraseologia e segnali discorsivi per studenti di italiano L2/LS*. Perugia, Guerra Edizioni.

Zamora Muñoz, Antonio Pablo (2005): "Fraseologia periferica e non solo". En Almela Perez, Ramón; Ramon, Trives y Gerd, Wotjak (ed.): *Fraseología contrastiva. Con ejemplos tomados del alemán, español, francés e italiano*. Murcia, Universidad de Murcia - Universität Leipzig, 65-80.

Zamora Muñoz, Pablo y Ramón, Almela Pérez (2005): "Multinomi: uno studio contrastivo italiano-spagnolo". *Rassegna Italiana di Linguistica Applicata*, XXXVII, 2-3, 183-199.

Zamora Muñoz, Pablo (2004): "La fraseología italiana y su enseñanza y aprendizaje". *Paremia*, 13, 125-136.

Zamora Muñoz, Pablo (2003): "Esquemas sintácticos fraseológicos pragmáticos". En Almela Pérez, Ramón (et al.) (ed.): *Homenaje al*

Profesor Estanislao Ramón Trives. Murcia, Universidad de Murcia, 825-836.

Zamora Muñoz, Pablo (1999a): "Unità fraseologiche pragmatiche in italiano". *Studi Italiani di Linguistica Teorica e Applicata*, XXVIII, 3, 547-556.

Zamora Muñoz, Pablo (1999b): "Otro intento de clasificación y delimitación de unidades fraseológicas italianas". En Ladrón de Guevara, Pedro Luís (et al.) (ed.): *Homenaje al Profesor Trigueros Cano*. Murcia, Ediciones Universidad de Murcia, 733-750.

Zamora Muñoz, Pablo (1999c): "Equivalencias lingüísticas de algunas frases idiomáticas pragmáticas en italiano". *Paremia*, 8, Madrid, 529-535.

Zamora Muñoz, Pablo (1997): *Spagnolo-Italiano. Espressioni idiomatiche e proverbi*. Milano, EGEA.

Zamora Muñoz, Pablo (1996): "Análisis contrastivo español-italiano de expresiones idiomáticas y refranes", *Paremia*, 5, 87-94.

Zingarelli, Nicola (2003): *Vocabulario della lingua italiana.* Bologna, Zanichelli.

Zuccalà Daniele (2018), "Fraseodidattica e Italiano LS. Studio di caso a Málaga (Spagna) e Quito (Ecuador)". En: Benucci, Elisabetta: *Fraseologia, Paremiologia e lessicografia III Convegno di PHRASIS*, Roma, ARACNE (Topoi 5).

Zuluaga, Alberto (1997): "Sobre fraseoloxismos e fenómenos colidantes". En Ferro Ruibal, Xesús; María, Jesús Anido Silvosa: *I Coloquio Galego de Fraseoloxía*, Santiago de Compostela, Xunta de Galicia,15-30.

Zuluaga, Alberto (1992): "Spanische Phraseologie". *Lexikon der RomanistischenLinguistik.* Tübingen, Niemeyer, 125-131.

Zuluaga, Alberto (1980): *Introducción al estudio de las expresiones fijas*. Frankfurt a. M./Bern, Peter Lang.

7.2. Índice de anexos (CD)

7.2.1. Archivo fraseológico_1.

Archivo fraseológico realizado tras la extracción manual de la fraseología a partir de *Il Nuovo Espresso.*

7.2.2. Archivo fraseológico_2.

Archivo fraseológico realizado tras la extracción manual de la fraseología a partir de *Il Nuovo Affresco Italiano.*

7.3. Resumen en italiano

Sunto in italiano

7.3.1. Contenuti e obiettivi della tesi

La presente tesi di dottorato «Indicatori fraseodidattici: una proposta di sistematizzazione della fraseologia italiana secondo il Quadro comune di riferimento per la conoscenza delle lingue (QCER)» è stata elaborata per rispondere alla necessità di integrare attraverso proposte empiriche la fraseologia nelle classi di italiano come lingua straniera (ILS). Per raggiungere detto obiettivo generale, la tesi di dottorato si suddivide in sette capitoli.

Il Capitolo I illustra le caratteristiche dello studio, fornisce una descrizione generale del contenuto dei capitoli che compongono la tesi e, infine, definisce le ipotesi e individua gli obiettivi che la ricerca si propone.

Con l'obiettivo di approfondire l'oggetto di studio nel Capitolo II, viene presentata una dettagliata analisi documentaria che, da un lato, permette di inquadrare la ricerca nello spazio e nel tempo e, dall'altro, di analizzare criticamente la letteratura precedente a questo lavoro. Si analizzano numerose fonti primarie e secondarie che hanno dato origine a tassonomie della fraseologia non solo italiana ma anche spagnola, in quanto le classificazioni di quest'ultima lingua hanno avuto una notevole influenza su quelle italiane negli ultimi decenni.

Nel capitolo seguente (Capitolo III), lo studio si concentra sullo stato dell'arte del concetto di fraseologia, inquadrandone le origini e definendone il concetto. Il QCER viene analizzato da un punto di vista puramente fraseologico, con l'obiettivo di osservare come viene trattata la fraseologia e, di conseguenza, quali sono le linee guida offerte agli insegnanti che desiderano lavorare con le unità fraseologiche (UF) nei diversi livelli del QCER.

Successivamente si studiano numerose proposte glottodidattiche volte a guidare gli insegnanti nel decidere in quale momento del ciclo formativo dello studente inserire la fraseologia nelle lezioni di lingua straniera e si analizzano le principali tendenze che difendono o osteggiano l'introduzione della fraseologia a partire dai livelli più elementari del QCER (A1, A2). Vengono inoltre presentate le principali difficoltà incontrate dagli insegnanti nell'insegnamento della fraseologia nelle classi di lingua straniera e, oltre ad analizzare le fonti primarie e secondarie, si menziona lo studio di caso che il dottorando ha presentato come tesi di master, tesi in cui si erano studiate tali difficoltà.

Successivamente si analizza il trattamento della didattica della fraseologia nei libri di testo di italiano e di spagnolo come lingue straniere e si presenta un'analisi approfondita di come la fraseologia viene trattata in due manuali di ILE: uno in cui viene trattata in modo pragmatico e scientifico, e un altro che rappresenta un modello incompleto e soggettivo. Infine, l'ultima sezione del capitolo è dedicata all'analisi della fraseologia italiana e spagnola nei dizionari: monolingui spagnoli, monolingui italiani e bilingui italiano-spagnolo.

Il capitolo IV si concentra su quelli che nella tesi sono stati definiti come «Indicatori Fraseodidattici» (IF), ovvero un insieme di strumenti diagnostici, sia qualitativi che quantitativi, che permettono la didattizzazione delle unità fraseologiche e la cui applicazione consente di distribuire la fraseologia in funzione dei livelli del QCER da un punto di vista didattico. In primo luogo, vengono identificati i principali IF attraverso un'attenta analisi delle proposte che precedono questo studio e, in secondo luogo, si analizzano le principali proposte metodologiche per l'applicazione di questi indicatori nelle classi di lingua straniera.

Il capitolo V definisce il disegno metodologico utilizzato in questo lavoro di dottorato. Vengono presentate le origini del modello sequenziale esplorativo proposto da Creswell (2003, 2007, 2009, 2014) e viene definita

una metodologia qualitativa e quantitativa che cerca di ottenere il maggior numero di informazioni sull'oggetto di studio attraverso diverse tecniche applicate in sequenza. Seguendo questo modello, la metodologia di questa ricerca è stata suddivisa in quattro fasi.

In seguito alla presentazione dell'approfondita rassegna documentaria (fase 1) precedentemente descritta, si procede con l'estrazione della fraseologia da due serie complete di manuali ILE (fase 2); a tal fine, si descrivono innanzitutto le caratteristiche del corpus selezionato: due serie complete (dal livello A1 al C2) di due metodi recenti di riferimento per l'insegnamento dell'italiano come lingua straniera (ILE). Per verificare la validità del processo di estrazione automatica della fraseologia, viene realizzato un pre-test utilizzando tre strumenti di estrazione automatica (SynchroTerm, Sketch Engine e il Term Extraction Module di IATE, il database terminologico dell'Unione Europea). Poiché i risultati si sono dimostrati insoddisfacenti, si è deciso di procedere con l'estrazione manuale della fraseologia. Il metodo utilizzato viene in seguito definito sulla base di standard internazionali per l'annotazione della fraseologia in un archivio fraseologico[263] elaborato *ad hoc* in formato elettronico. La presentazione quantitativa della fraseologia estratta completerà l'ultima parte della seconda fase della metodologia.

Nella successiva fase 3, viene presentato un modello sequenziale per identificare la frequenza delle unità fraseologiche precedentemente estratte (*Refence corpus*) in un corpus generale della lingua (*Focus Corpus*). A tal fine, viene presentata la proposta di creazione di uno strumento che si realizzerà in collaborazione con lo *IATE Support & Development Team*, entità interistituzionale con cui il dottorando ha

[263] Nel presente studio, il termine archivio fraseologico è utilizzato con un significato assimilabile a una raccolta/elenco elettronico (formato .XLSX) di UF raccolte e organizzate sulla base di criteri comuni. Il termine archivio fraseologico è preso in prestito dalla terminologia e definito «un insieme di record terminologici che sono logicamente collegati in un database utilizzando lo stesso modello di presentazione, applicando le stesse regole di inserimento e utilizzando un unico modello preciso di interrogazione» (Pavel e Nolet, 2001: 127).

lavorato presso il Centro di traduzione dell'Unione Europea. Nella sezione seguente si presenta un'analisi strategica dei rischi e dei punti di forza di tale strumento, con l'idea di limitarne gli aspetti negativi e rafforzarne le opportunità.

Questo strumento, basato sui principi dell'elaborazione del linguaggio naturale (NLP), consentirà di accedere all'ultima fase del modello metodologico (fase 4), ossia alla proposta di applicazione degli indicatori fraseodidattici alle unità fraseologiche precedentemente estratte e catalogate. In questo senso, si propone di applicare i risultati della fase precedente ai diversi livelli del QCER e di definire il livello gerarchico occupato da ciascuna unità fraseologica estratta all'interno di un livello di riferimento del QCER.

Il capitolo VI presenta le conclusioni e le future linee di ricerca mentre i riferimenti bibliografici e i due archivi fraseologici, risultato dell'estrazione e dell'annotazione della fraseologia, chiuderanno la tesi di dottorato con il capitolo VII.

7.3.2. Metodologia seguita

Con il fine di studiare quelli che in questo lavoro sono stati definiti come "Indicatori Fraseodidattici" (IF) applicabili alla didattica della fraseologia nelle classi di italiano come lingua straniera (ILE), si è utilizzato un disegno metodologico qualitativo e quantitativo. In particolare, il modello sequenziale esplorativo (Cresswell, 2003, 2007, 2009, 2014).

Questo modello metodologico proposto da John Cresswell cerca di ottenere il maggior numero di informazioni possibili sull'oggetto di studio, in questo caso l'insegnamento delle unità fraseologiche nelle classi di italiano come lingua straniera, attraverso diverse tecniche applicate in sequenza. Il disegno metodologico avanza grazie all'applicazione sequenziale di diversi metodi: i risultati ottenuti nella prima fase permettono di accedere alla seconda fase, questa a sua volta consentirà

di procedere con la terza e così via fino all'ultima fase che completerà la metodologia.

Il disegno multi-metodo proposto dal professore della New Jersey City University è stato adattato alle esigenze del presente studio attraverso quattro diverse fasi (vedi Figura 19). Sono stati utilizzati metodi quantitativi e qualitativi con l'obiettivo di analizzare la complessa realtà delle unità fraseologiche applicate alla didattica nelle classi di italiano come lingua straniera. In breve, la metodologia del presente lavoro si basa su una logica induttiva che permette di progettare una proposta cumulativa basata su quattro fasi consecutive:

1. Revisione della letteratura;

2. Estrazione e annotazione della fraseologia;

3. Proposta per l'analisi della frequenza della fraseologia nel *Focus corpus*;

4. Proposta di applicazione degli indicatori fraseologici al Reference corpus[264].

Figura 19. Disegno metodologico sequenziale esplorativo utilizzato.



Fonte: Elaborazione propria.

La prima fase ha comportato una revisione documentale delle fonti primarie e secondarie. A tal fine, sono state analizzate criticamente le principali proposte metodologiche utilizzate per l'insegnamento della fraseologia italiana e spagnola, con l'obiettivo di presentare olisticamente

[264] Scelta dei termini *Focus corpus* e *Refence corpus* in base alla terminologia utilizzata da Sketch Engine (vedi sezioni 5.3.1. e 5.3.2.).

lo stato della questione e porre le basi per future linee di ricerca. È stato inoltre condotto uno studio dettagliato delle principali proposte di lavoro per trattare la fraseologia nelle classi di italiano e spagnolo come lingua straniera. In primo luogo, è stata presentata un'analisi critica degli indicatori fraseodidattici più produttivi per l'organizzazione del materiale fraseologico, analizzando esclusivamente le proposte basate su criteri empirici e tralasciando quelle che presentano diversi gradi di soggettività, cioè quelle che affondano le radici esclusivamente nell'esperienza personale del loro autore. È da notare che, per la prima volta, vengono sistematizzati in un'unica opera i principali indicatori fraseologici per la stesura di una lista di UF applicate alla didattica dell'italiano come lingua straniera.

Dopo la revisione documentale, la seconda fase della metodologia prevede l'estrazione manuale della fraseologia da due serie complete (A1-C2) di manuali di riconosciuto prestigio nel campo dell'insegnamento dell'italiano come lingua straniera: *Nuovo Espresso* e *Il Nuovo Affresco Italiano*. Si è scelto di realizzare un'estrazione manuale piuttosto che una automatica soprattutto a causa delle limitazioni ancora presenti negli strumenti di estrazione terminologica. È stato infatti effettuato un pre-test utilizzando tre strumenti di estrazione automatica, ovvero: Sketch Engine di Lexical Computing Limited, SynchoTerm di Terminotix e, infine, il recente Term Extraction Module (TEM) di IATE, Interactive Terminology for Europe. Tuttavia, i risultati sono stati insoddisfacenti per diverse ragioni definite nella sezione 5.2.2.1.

Dopo aver identificato ed estratto la fraseologia rilevante per il presente studio, è stato creato un archivio o file fraseologico *ad hoc* in formato digitale in cui sono state annotate e trasferite le UF. Con il termine archivio fraseologico si identifa un file online (formato .XLSX) che presenta un elenco di UF ordinate in base a criteri comuni (cfr. sezione 1.5.). La creazione di questo strumento si è basata sugli standard

dell'*Organizzazione Internazionale per la Standardizzazione* (ISO)[265] e in particolare sulla ISO 704: 2000, *Terminology work - Principles and methods* e ISO 1087-2: 2000, *Terminology work - Vocabulary - Part 2: Computer applications*. Detti standard fondamentali per il lavoro terminologico sono stati applicati all'area della fraseologia.

La terza fase prevede il confronto delle unità fraseologiche estratte (*Reference corpus*) con un corpus di lingua italiana generale (*Focus corpus*) con l'obiettivo principale di analizzare il comportamento della frequenza sulle UF oggetto di studio. A tal fine, il presente studio propone la creazione di un software creato in collaborazione con lo *IATE Support and Development Team* che permetterà di analizzare le unità estratte manualmente in un corpus generale della lingua italiana composto da oltre 4,9 miliardi di parole, e di restituire il livello di frequenza assoluto di ogni unità fraseologica nella lingua generale.

Per analizzare la frequenza delle unità fraseologiche estratte si è deciso di effettuare un test preliminare con lo strumento Sketch Engine. L'applicazione *online* offre diversi corpora di grandi dimensioni e, per la lingua italiana, l'*Italian Web 2020* (itTenTen20[266]) è quello con maggiori dimensioni. La versione più recente del corpus itTenTen è composta da 12 miliardi di parole e i testi che compongono il corpus riflettono una lingua italiana attuale in quanto sono stati scaricati dal web tra novembre e dicembre 2019 e dicembre 2020. itTenTen fa parte della famiglia di corpora TenTen, un insieme di oltre 30 corpora compilati via web, ciascuno con una dimensione di oltre 10 miliardi di parole.

[265] L'ISO è un'organizzazione internazionale indipendente e non governativa fondata nel 1946 con sede a Ginevra. L'ISO ha 164 organismi di normazione nazionali e il suo scopo principale è lo sviluppo di norme internazionali, cioè di norme basate sul consenso internazionale e rilevanti per il mercato. Per ulteriori informazioni, consultare il sito https://www.iso.org/standards.html.

[266] Per ulteriori informazioni, vedere < https://www.sketchengine.eu/ittenten-italian-corpus/?gclid=Cj0KCQjw7ZL6BRCmARIsAH6XFDICVF2Zm5L0V3qeQSW7ctaQriZbNEHeMKM0NNzr4Zy45PQEla4zIvUaAqjEEALw_wcB>.

Sketch Engine permette di confrontare due corpora: il proprio (chiamato *Reference corpus*) e il corpus itTenTen20 (definito come *Focus corpus*). Tuttavia, dopo diversi tentativi falliti poiché l'applicazione restituiva risultati del tutto insoddisfacenti in termini di fraseologia, si è deciso di contattare il team di Sketch Engine per chiedere una spiegazione e una soluzione. Dopo uno scambio di e-mail, il team di supporto ha confermato che non era possibile analizzare automaticamente la frequenza di numerose unità fraseologiche, ma che le UF dovevano essere aggiunte manualmente, una per una, nel sistema e ricercate nel *Focus corpus*. Si tratta ovviamente di una soluzione che richiede molto tempo e che è stata quindi esclusa a priori. In questo senso, si osservano i limiti evidenziati da Hidalgo-Ternero e Corpas Pastor (2020) quando affermano che «vent'anni dopo la pubblicazione seminale di Sag et al. (2002), le espressioni multiverbali (MWE) sono ancora un grattacapo per i sistemi di elaborazione del linguaggio naturale»[267].

Data l'impossibilità di utilizzare uno strumento online esistente[268], nasce l'esigenza di sviluppare uno strumento *ad hoc* per identificare la frequenza di un gran numero di unità fraseologiche in un corpus generale della lingua. Per questa proposta ci si è avvalsi dell'esperienza negli strumenti di elaborazione del linguaggio naturale (NLP), e in particolare nell'estrazione terminologica, del gruppo IATE con cui il dottorando ha lavorato per otto mesi presso il Centro di traduzione dell'Unione europea come *business analyst*, tester e coordinatore della piattaforma EurTerm[269].

[267] «Twenty years after Sag et al.'s publication seminal publication, multiword expressions (MWEs) are still a pain in the neck for natural processing systems». Traduzione propria.

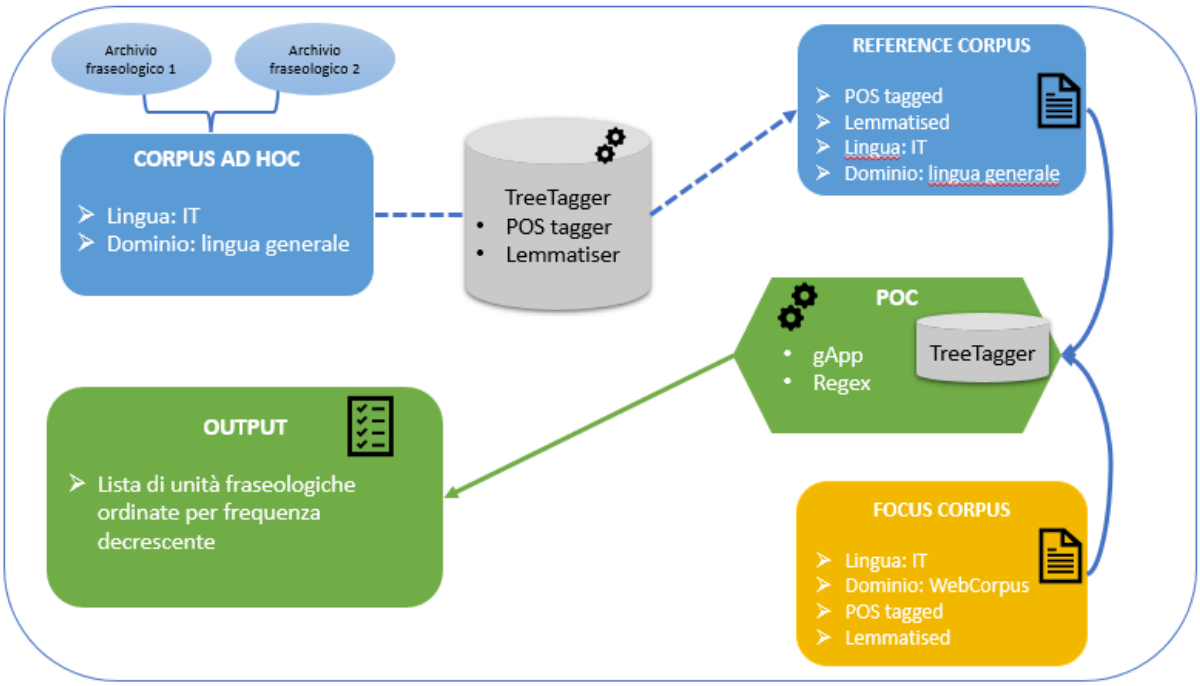
[268] Purtroppo, le risorse di elaborazione del linguaggio naturale (NLP) dedicate alla fraseologia monolingue sono ancora molto scarse, una dimostrazione pratica del fatto che «la fraseologia e l'NLP non sono state veramente in contatto, nonostante si occupino di fenomeni molto simili: unità fraseologiche ed espressioni multi-verbali» (Corpas Pastor, 2013:368, Traduzione propria.).

[269] EurTerm è il portale terminologico interistituzionale creato e gestito dal gruppo di gestione IATE. Il portale è stato creato con l'obiettivo di promuovere la cooperazione terminologica tra le istituzioni dell'UE e contiene risorse di tutte le istituzioni partecipanti. Per ulteriori informazioni su EurTerm si veda <https://termcoord.eu/intra-eu/eurterm/>.

Va notato che questa terza fase della metodologia è presentata come una proposta, poiché è stato impossibile iniziare a scrivere il codice dello strumento nei tempi stabiliti per la difesa della tesi. Dopo diversi incontri con il gruppo IATE vennero gettate le basi per la sua definizione, ma a causa delle priorità urgenti definite dai partner interistituzionali, non è stato ancora possibile stabilire una data esatta per avviare lo sviluppo informatico dello strumento. Inoltre, la situazione pandemica causata dalla COVID19, il telelavoro forzato e le assenze del personale hanno complicato ulteriormente la situazione.

La figura 20 rappresenta una definizione semplificata del processo che permetterà di identificare la frequenza della fraseologia attraverso uno strumento *ad hoc*.

Figura 20. Modello utilizzato per identificare la frequenza delle UF.



Fonte: elaborazione propria.

Infine, la quarta e ultima fase della presente metodologia prevede l'applicazione degli indicatori fraseologici alla fraseologia estratta e precedentemente confrontata con il corpus generale della lingua

italiana. In questo studio si propone di classificare le UF estratte dai manuali ILE (seconda fase della metodologia) stabilendo un livello gerarchico in base alla loro frequenza per stabilire in quale livello del QCER insegnare/apprendere una certa unità fraseologica. I blocchi per la suddivisione della fraseologia sono proposti come segue: a ogni rango corrisponde un livello QCER che viene aumentato di 1,9 punti in base alla frequenza a cui appartiene un'unità, ad eccezione dei ranghi 5 e 6 (C1 e C2) che vengono aumentati di un punto ciascuno. Vale a dire:

Tabella 5. Gruppi di frequenza in base al QCER

Gruppo	**Frequenza**	**Livello del QCER**
Gruppo 1	tra 10 e 8,1	A1
Gruppo 2	tra 8,0 e 6,1	A2
Gruppo 3	tra 6,0 e 4,1	B1
Gruppo 4	tra 4,0 e 2,1	B2
Gruppo 5	tra 2,0 e 1,1	C1
Gruppo 6	tra 1,0 e 0,1	C2

Fonte: Elaborazione propria.

I gradi 5 e 6, che corrispondono ai livelli C1 e C2 del QCER, sono aumentati di un solo punto per due motivi principali: in primo luogo, come spiegato nel capitolo 3.3., è necessario insegnare un numero maggiore di UF a partire dai livelli più elementari del QCER. In secondo luogo, trattandosi di unità poco frequenti nella lingua, si ipotizza un numero maggiore di UF in questi due intervalli, da cui la necessità pratica di ridurne la frequenza a un punto.

La presente proposta consente di assegnare a ogni unità fraseologica un livello di riferimento del QCER, in modo che sia gli insegnanti che i professionisti coinvolti nella preparazione di libri di testo di italiano come

lingua straniera possano utilizzare una prima proposta empirica nella creazione dei materiali per la classe.

7.3.3. Rischi e limiti della metodologia

Lavorare con *Big Data*[270] presenta grandi vantaggi e punti di forza, ma allo stesso tempo limiti che vale la pena sottolineare. Il principale limite derivante dalla metodologia impiegata risiede nello strumento informatico che si utilizzerà per estrarre la frequenza delle UF compilate da un corpus generale della lingua italiana. Detto strumento presenta una serie di rischi che si descrivono di seguito.

Il primo rischio è di natura puramente tecnica e riguarda le prestazioni informatiche dell'applicazione. La caratteristica principale di entrambi gli input, e del *Focus corpus* in particolare, è che presentano un insieme di dati così voluminosi da rallentarne in modo significativo l'analisi e la generazione dell'output. Inoltre, l'atto stesso di caricare file di grandi dimensioni nell'applicazione potrebbe influire negativamente sulle prestazioni del software. Pertanto, la velocità di calcolo, il tempo di risposta del software per reagire all'input (*response time*) e le prestazioni dell'applicazione dovranno essere monitorate sia nella fase di sviluppo che in quella di test dello strumento e dovranno essere gestite come un rischio di progetto, al fine di trovare la migliore soluzione pratica per evitarlo o limitarlo.

La seconda limitazione è di natura puramente linguistica ed è inerente alla natura semi-fissa della fraseologia. Sia nel discorso parlato che in quello scritto, le unità fraseologiche non appaiono sempre come una sequenza contigua di parole, ma in alcune unità è possibile intercalare nuovo lessico senza sacrificare il significato compositivo o idiomatico dell'unità. Nella frase «mio figlio sembra aver messo leggermente la testa

[270] Per *Big data* si intende un insieme di dati informatici così voluminosi che un software di elaborazione tradizionale non sarebbe in grado di gestire. Grazie alle nuove tecnologie, questi enormi volumi di dati possono essere utilizzati per risolvere problemi teorici ed empirici che prima non potevano essere affrontati. Per maggiori informazioni si veda Sagiroglu e Sinanc (2013).

a posto», l'unità fraseologica «*mettere la testa a posto*» viene interrotta dall'avverbio «*leggermente*» senza che ciò influisca sul suo significato compositivo. Tuttavia, uno strumento di rilevamento automatico potrebbe avere difficoltà a individuare detta UF, il che influenzerebbe negativamente i risultati.

Numerosi studi hanno analizzato questo problema e proposto soluzioni[271]. In particolare, il lavoro di Hidalgo-Ternero e Corpas Pastor ha compiuto progressi significativi nell'individuazione automatica di unità fraseologiche discontinue. L'algoritmo ReGap, utilizzato nell'applicazione online gApp, si basa su uno studio precedente dei modelli presenti nel corpus e consente di estrarre e ordinare la fraseologia individuata in un testo. Nelle parole degli autori, lo strumento gApp è «un sistema di pre-elaborazione del testo basato su Python per l'identificazione e la conversione automatica di espressioni fraseologiche discontinue nella loro forma continua, al fine di migliorare la traduzione automatica neurale (NMT)» (Hidalgo-Ternero e Corpas Pastor, 2020: 61).[272]

Il suo algoritmo, programmato in Phyton, comprende anche *espressioni* regolari (*Regex*) create *ad hoc* per ogni unità fraseologica con l'obiettivo di perfezionare l'individuazione di somatismi non consecutivi[273] . Questo strumento rappresenta un importante passo avanti nello stato dell'arte del rilevamento automatico della fraseologia, in quanto è in grado di rilevare le UF discontinue in un testo e, d'altro canto, le riordina nella loro forma continua.

L'algoritmo è stato implementato per diverse lingue[274] e ha un enorme potenziale per il presente studio, in quanto permetterebbe di individuare

[271] Il lavoro di Hidalgo-Ternero e Corpas Pastor (2020) fornisce un attento e dettagliato stato dell'arte in termini di rilevamento automatico di unità fraseologiche discontinue.
[272] «Un sistema de preprocesamiento de textos basado en Python para la identificación y conversión automática de expresiones fraseológicas discontinuas a su forma continua con el fin de mejorar la traducción automática neural (NMT)». Traduzione propria.
[273] L'obiettivo finale dello studio era quello di analizzare come tale algoritmo potesse migliorare le prestazioni dei sistemi di traduzione automatica neurale (NMT).
[274] Ovvero: spagnolo, inglese, francese, tedesco, italiano e cinese.

le unità fraseologiche discontinue nel *Focus corpus* e di aumentare l'affidabilità dell'output. Va notato che non è possibile implementare tale algoritmo direttamente nello strumento, ma sarebbe necessario uno studio preliminare di ogni UF che compone il *Refence corpus*. Tuttavia, date le sue potenzialità, l'integrazione di tale algoritmo nel codice dello strumento potrebbe limitare il rischio derivato dalla discontinuità della fraseologia e preservarne i risultati.

Da un punto di vista semantico, un altro rischio da considerare causato dall'analisi automatico della frequenza di certe unità in un corpus di grandi dimensioni è la collisione omonimica (Penadés Martínez, 2015), concetto descritto in dettaglio nelle sezioni precedenti (cfr. 4.1.8.). In breve, la collisione omonimica corrisponde all'omonimia che si verifica tra una locuzione e una combinazione di parole non fraseologica. Nella frase «Daniele ha tagliato la corda», l'unità *tagliare la corda* può essere una locuzione idiomatica (andarsene precipitosamente per mettersi in salvo o per scansare un dovere) o una combinazione libera di parole a partire dal verbo *tagliare* e dal sostantivo *corda*. Un'analisi umana realizzata da un madrelingua sarebbe in grado di identificare il significato compositivo, ma uno strumento informatico potrebbe confondere facilmente l'unità fraseologica con la combinazione libera di parole e viceversa.

Finora non è stata ancora trovata una soluzione definitiva per l'individuazione automatica dei casi di collisione tra omonimi, per cui l'analisi manuale delle linee di concordanza di ogni UF sarebbe sicuramente la più appropriata. Tuttavia, data l'enorme quantità di dati presenti nel *Focus corpus*, un'analisi manuale potrebbe essere a priori un grande sforzo che non restituirebbe un beneficio adeguato. Ad esempio, per quanto riguarda i somatismi, Hidalgo-Ternero e Corpas Pastor (2020) hanno dimostrato che il 50% delle unità trovate[275] corrispondeva a una combinazione libera di parole (*irrelevant result*, per usare la terminologia

[275] Gli autori hanno analizzato manualmente un totale di 560 casi di quattro somatismi e hanno trovato 70 risultati irrilevanti per ciascuno di essi, cioè casi di collisione di omonimi.

degli autori). Se si dimostrasse che una percentuale simile si applica anche ad altri tipi di unità diverse dai somatismi, si dovrebbe senza dubbio trovare una soluzione, manuale o automatica, per limitare l'impatto della collisione di omonimi sui risultati. A questo proposito, sarebbe necessario approfondire l'argomento e studiare il rischio nel dettaglio, anche in base al tipo di unità fraseologica.

Oltre ai rischi previamente menzionati, lo strumento presenta anche punti di forza e vantaggi come la flessibilità dello strumento che è senza dubbio uno delle sue maggiori virtú. Come visto in precedenza, l'interfaccia grafica del software consentirà di caricare due input, il che significa che cambiando il *Focus corpus* sarà possibile analizzare come variano i risultati forniti e quindi perfezionare lo studio.

Inoltre si tratta di uno strumento monolingue che non si concentra su nessuna lingua in particolare. Può essere quindi utilizzato per svolgere lo stesso studio ma su una lingua differente, il che porterà un grande valore aggiunto alla fraseologia e alla fraseodidattica di altre lingue. Va notato che il modello di lemmatizzazione definito sopra è più adatto per lavorare con lingue analitiche come l'italiano, mentre per le lingue agglutinanti si preferiscono di solito i processi *di stemming*, cioè il processo di rimozione automatica degli affissi (suffissi, prefissi, ecc.) da una parola, lasciando solo la radice o lemma. A tal fine, è necessario analizzare i risultati ottenuti con le lingue agglutinanti e, se necessario, elaborare i corpora con il processo di *stemming* anziché con la lemmatizzazione.

7.3.4. Considerazioni finali

Il presente lavoro di dottorato ha presentato un'approfondita rassegna e analisi delle principali teorie e classificazioni della fraseologia e degli studi fraseodidattici, con particolare attenzione alla lingua italiana e spagnola. La prospettiva critica utilizzata in questa analisi ha permesso di osservare che la fraseologia spagnola e italiana non hanno ricevuto lo stesso interesse accademico nel corso del tempo. Tuttavia, è stato possibile

indagare come la fraseologia italiana, pur avendo ricevuto un interesse minore rispetto a quella spagnola, abbia registrato una proliferazione di studi soprattutto negli ultimi decenni.

È stato anche osservato che la fraseologia spagnola è servita, e serve tuttora, come modello per quella italiana e che lo stesso fenomeno di globalizzazione e di scambio di conoscenze a livello mondiale favorisce un clima di analisi e di studio della fraseologia italiana. Si è dimostrato che esiste un numero significativo di accademici - la cui lingua materna è lo spagnolo – che di dedicano alla fraseologia italiana e le cui analisi sono servite a favorire lo sviluppo degli studi fraseologici italiani negli ultimi decenni.

Numerose tassonomie sono state proposte da italianisti per classificare la fraseologia italiana (cfr. Cap. II), anche se principalmente da una prospettiva lessicografica e spesso limitati ad alcuni tipi di fraseologismi specifici. In particolare, si è notato che la sfera delle locuzioni è quella che ha ricevuto il maggiore interesse nella fraseologia italiana, fatto che trova dimostrazione pratica anche nei manuali analizzati, dove la maggior parte delle unità estratte corrisponde alla sfera delle locuzioni. Inoltre, è stato osservato un certo «caos terminologico» (Mellado Blanco, 2010: 358): la mancanza di una terminologia accettata dalla maggioranza degli italianisti ne ha generata una instabile e variegata. Quanto detto si è riscontrato sia negli studi teorici che nei manuali di italiano come lingua straniera.

D'altra parte, lo studio ha approfondito il concetto di fraseodidattica, calco del termine *phraseodidactik* coniato da Kühn (1987), e ha adottato la definizione proposta da González Rey (2012) che comprende una concezione linguistica, semantica e pragmatica o, in altre parole, culturale dell'insegnamento della fraseologia. Per questa ragione, si è utilizzato il termine fraseodidattica per indicare la scienza dell'insegnamento e dell'apprendimento delle unità fraseologiche, come

definite nella classificazione di Corpas Pastor (1996), da un punto di vista grammaticale, pragmatico e semantico.

Studiando in dettaglio come la fraseologia italiana è stata applicata alla didattica dell'ILE, il primo obiettivo generale di questo lavoro è stato raggiunto e la prima ipotesi del presente lavoro di dottorato è stata corroborata, dimostrando che la maggior parte delle proposte di applicazione degli indicatori fraseodidattici alla fraseologia italiana soffre di un certo grado di soggettività e che le raccolte di UF applicate alla didattica della fraseologia non sono compilate in modo empirico, ma si basano in gran parte sulla mera esperienza soggettiva di chi le propone.

Per ovviare a questa mancanza sono state analizzate diverse proposte accademiche per applicare gli IF all'insegnamento della fraseologia in classe, il che ha permesso di raggiungere il secondo obiettivo principale di questo studio, osservando che la frequenza è uno degli indicatori fraseodidattici più rilevanti ed efficaci per inquadrare una UF in uno specifico livello del QCER. Tutto ciò ha permesso di avvalorare la seconda ipotesi di questo studio dottorale.

L'uso di una metodologia sequenziale esplorativa (Creswell, 2003, 2007, 2009, 2014) ha consentito di compilare un corpus composto da unità fraseologiche estratte da due serie complete di manuali ILE. Il pre-test effettuato durante l'estrazione ha portato alla verifica della terza e ultima ipotesi, secondo la quale l'estrazione manuale della fraseologia per questa tipologia di studio fornisce risultati più affidabili rispetto all'estrazione automatica, sia dal punto di vista quantitativo che qualitativo.

Infine, la proposta di un modello per identificare la frequenza di un corpus composto da unità fraseologiche ha permesso di raggiungere il terzo obiettivo principale di questo studio di dottorato. Questo modello, basato sulla creazione di uno strumento informatico, permetterà di confrontare un corpus composto da UF grammaticalmente etichettate e

lemmatizzate (*Reference corpus*) con un corpus generale della lingua italiana (*Focus corpus*). L'algoritmo terrà conto dei rischi legati al rilevamento automatico di unità fraseologiche discontinue (Hidalgo-Ternero e Corpas Pastor 2020), nonché del fenomeno della *colisión homonímica* (Penadés Martínez, 2015) inteso come l'omonimia che si produce tra una locuzione e una combinazione libera di parole.

7.3.5. Future linee di ricerca

Tenendo conto delle considerazioni finali e delle limitazioni direttamente connesse alla natura di uno studio di dottorato, di seguito vengono illustrate le future linee di ricerca relative all'oggetto di studio.

L'obiettivo a breve termine è sviluppare il software per estrarre la frequenza delle UF compilate. In collaborazione con il team di IATE, verrà definita in dettaglio l'architettura dell'applicazione. Aspetti quali l'architettura del sistema (back-end), l'interfaccia grafica (GUI), l'implementazione dell'algoritmo della macchina di supporto vettoriale (classificatore SVM), il *fine-tuning*, l'integrazione di Regex nell'algoritmo, ecc., saranno analizzati e definiti in un progetto specifico.

La presente tesi di dottorato è di per sé un chiaro impegno nella continuità degli studi di fraseologia e fraseodidattica dell'italiano come lingua straniera. È il primo passo verso un modello più complesso di fraseodidattica che tenga conto di una gamma più ampia di indicatori fraseologici. Detta proposta sarà infatti perfezionata applicando nuovi IF all'elenco delle UF. L'obiettivo è di prendere in considerazione quegli IF che, pur non essendo così significativi come la frequenza delle unità nella lingua generale, permetteranno di perfezionare la distribuzione della fraseologia italiana nei diversi livelli del QCER.

Come primo indicatore fraseodattico si prenderà in considerazione quello denominato «Idiomaticità ed equivalenza», essendosi rivelato uno degli IF più rilevanti per la comunità scientifica. Da una parte, l'idiomaticità

fraseologica, una proprietà non necessaria delle UF per la quale il significato dell'unità non è deducibile dai singoli componenti, suscita grande curiosità negli studenti e facilita la memorizzazione delle unità proposte. D'altra parte, l'equivalenza di UF tra lingue diverse, in particolare quelle affini, facilita la loro fissazione nel lessico mentale di un parlante, per cui potrebbe essere sfruttata in classi omogenee composte da studenti che condividono la stessa lingua madre.

Come obiettivo a lungo termine, si intende sviluppare materiali didattici per l'italiano come lingua straniera basati sulla metodologia presentata. Detti materiali rappresenterebbero un passo avanti per l'ILE, in quanto basati su un corpus compilato da manuali di riferimento. Nell'elaborazione di questi libri di testo si utilizzeranno i preziosi contributi del progetto GEPHRI (vedi sezione 4.3.4.) e si proporrà di collaborare con il gruppo di ricerca che guida questo progetto per offrire un valore aggiunto all'insegnamento e all'apprendimento dell'italiano come lingua straniera.

Infine, l'obiettivo è quello di applicare questa metodologia ad altre lingue romanze e, in particolare, allo spagnolo. Poiché lo strumento monolingue non si basa su nessuna lingua in particolare, ma è in grado di estrarre la frequenza di una UF dai corpora, sarà sufficiente compilare un nuovo Reference corpus composto da fraseologia spagnola e contrastarlo con un Focus corpus della lingua spagnola generale. In questo modo, sarà possibile avere una lista di UF spagnole distribuite in funzione del QCER, aggiungendo valore alla già vasta letteratura fraseologica spagnola e fornendo all'insegnante di spagnolo come lingua straniera un prezioso strumento di lavoro.

7.3.6. Conclusioni

Per quanto affermato in precedenza, si può pertanto concludere che:

1. La fraseodidattica dell'italiano merita una maggiore attenzione da parte degli italianisti, poiché la didattica dell'italiano come lingua

straniera non può e non deve omettere la fraseologia dai suoi programmi e manuali.

2. È necessario insegnare fraseologia a tutti i livelli del QCER, dal più elementare (A1) al più avanzato (C2), e stabilire a priori quali UF insegnare a ciascun livello. Sebbene esista una corrente teorica che ritiene opportuno insegnare la fraseologia a partire dai livelli intermedi del QCER (il QCER stesso sembra incoraggiare questa posizione), un'analisi dei libri di testo ha dimostrato che nella pratica un numero significativo di unità fraseologiche viene insegnato a partire dal livello A1.

3. L'ambito delle locuzioni è quello che ha ricevuto maggiore attenzione nell'insegnamento della fraseologia italiana come lingua straniera. Questa tendenza è stata confermata sia dallo stato dell'arte che dall'analisi dei libri di testo ILE analizzati e che costituiscono il *Reference corpus del* presente lavoro.

4. L'insegnamento/apprendimento delle unità fraseologiche nei diversi livelli del QCER deve seguire un modello empirico che eviti la soggettività. Il presente studio è una proposta di sistematizzazione e non si esclude un suo futuro miglioramento, includendo ulteriori indicatori fraseodidattici che permettano di perfezionare l'assegnazione della fraseologia ai livelli QCER.

5. L'estrazione automatica e il *POS tagging* della fraseologia a partire da un dato corpus presenta alcuni limiti. Tuttavia, sono stati fatti molti passi avanti in questa direzione e questi progressi delineano una situazione promettente per il futuro della fraseologia nella linguistica dei corpora e computazionale.

6. L'uso di uno strumento informatico per studiare automaticamente la fraseologia presenta dei rischi che devono essere studiati in dettaglio per garantire che i risultati non vengano alterati. Tuttavia, è altamente

auspicabile lo sviluppo di uno strumento di NLP per studiare la fraseologia in termini di frequenza e assegnarla ai diversi livelli del QCER.

7. Proponendo un metodo per distribuire la fraseologia italiana in ciascun livello del QCER, questo lavoro propone di colmare una lacuna nella fraseodidattica dell'italiano come lingua straniera, rendendo il presente studio di dottorato necessario e pertinente.