

LEXICOGRAFÍA ESPECIALIZADA Y LENGUAJES DE ESPECIALIDAD: FUNDAMENTOS TEÓRICOS Y METODOLÓGICOS PARA LA ELABORACIÓN DE DICCIONARIOS ESPECIALIZADOS.

SPECIALISED LEXICOGRAPHY AND SPECIALISED LANGUAGES:
THEORETICAL AND METHODOLOGICAL FOUNDATIONS FOR THE
ELABORATION OF SPECIALISED DICTIONARIES

NURIA EDO MARZÁ

Universitat Jaume I, Castellón, España
nedo@ang.uji.es

En el presente artículo se pretende mostrar de forma progresiva y secuencial pero integrada una aproximación teórico-práctica al desarrollo de diccionarios especializados desde un enfoque activo, centrado en el usuario y sus necesidades y acorde con las corrientes terminográficas actuales de tipo lingüístico-comunicativo, lo cual nos llevará a adoptar finalmente la etiqueta de “lexicografía especializada” para referirnos a los fundamentos teórico-metodológicos que guían o deberían guiar el desarrollo de estos diccionarios de especialidad.

Con este fin y desde una perspectiva teórica se han analizado, entre otros aspectos, los tres pilares básicos en que se sustenta la lexicografía especializada: (1) los lenguajes de especialidad, (2) la terminología, entendida aquí desde la *Teoría Comunicativa de la Terminología* (Cabré 1999) y (3) la lingüística de corpus como medio necesario para el análisis de los términos en uso. Esta perspectiva teórica de la lexicografía especializada se presenta a su vez complementada por unos fundamentos prácticos y una propuesta metodológica concreta para la elaboración de diccionarios activos de especialidad, mostrándose aquí como resultado el tipo de entrada lexicográfica a obtener, desarrollada siempre y necesariamente en base a unos usuarios y a unas necesidades comunicativas específicas.

Recibido
23/10/11
Aceptado
11/01/12

Palabras clave: lexicografía especializada, lenguajes de especialidad, terminología, Teoría Comunicativa de la Terminología, corpus, diccionario activo.

The present article aims to put forward in a progressive and sequential but also integrated manner a theoretical-practical approach for the development of specialised dictionaries from an active perspective focused on the user, on his/her needs and in accordance with current terminographical research, linguistic-communicative in nature. This will lead to the final adoption of the label “specialised lexicography” to make reference to the theoretical-methodological foundations that rule or should rule the development of such speciality dictionaries.

With this aim in mind and from a theoretical perspective, the three main pillars on which specialised lexicography is sustained have been analysed, namely, (1) specialised languages, (2) terminology, understood here from the *Communicative Theory of Terminology* (Cabré 1999) and (3) corpus linguistics as the necessary means for the analysis of terms in use. At the same time, the theoretical perspective is presented in conjunction with a series of practical foundations and a specific methodological proposal for the elaboration of specialised active dictionaries. The kind of results to be obtained have been illustrated in the present study with a sample, real lexicographical entry, always developed on the basis of prospective users and specific communicative needs.

Keywords: specialised lexicography, specialised languages, terminology, Communicative Theory of Terminology, corpus, active dictionary.

1. FUNDAMENTOS TEÓRICOS PARA LA ELABORACIÓN DE DICCIONARIOS ESPECIALIZADOS

El presente estudio pretende presentar de forma secuencial pero integrada una aproximación teórico-práctica al desarrollo de diccionarios especializados desde una perspectiva activa, comunicativa y acorde con las corrientes terminográficas actuales, lo que nos llevará a adoptar finalmente la etiqueta de “lexicografía especializada” para englobar el tipo de enfoque teórico-práctico descrito. Para ello, analizaremos los tres pilares básicos en que se sustenta nuestra visión de la lexicografía especializada (entendida aquí como una terminografía de amplio espectro, de base lingüística y comunicativa y que va más allá de la compartimentación estanca del conocimiento) y que son (1) los lenguajes de especialidad, (2) la terminología como eje vertebrador del discurso especializado y entendida desde la Teoría Comunicativa de la Terminología (TCT) (Cabré 1999) y, a su vez, desde la lingüística y (3) la lingüística de corpus

Este recorrido nos llevará a entender qué es la lexicografía especializada y cómo enmarca y determina el modo de desarrollar diccionarios

activos de especialidad destinados a un fin y a usuarios con necesidades concretas, conduciéndonos de forma natural a establecer una praxis lexicográfica con una fundamentación metodológica clara y sistemática (ver apartado 2), claramente apoyada en los recursos terminóticos existentes, para la elaboración de dicho tipo de diccionarios.

1.1. El lenguaje general, la lengua común y los lenguajes de especialidad

Cuando un terminógrafo inicia la tarea de elaborar un diccionario de especialidad o especializado, existen muchos aspectos, tanto teóricos como prácticos, que debe conjugar de manera efectiva para lograr el resultado esperado tanto por él mismo (o su grupo) como por el potencial usuario de la obra. Esto implica entender desde un primer momento que un diccionario especializado no es sólo un compendio de palabras técnicas traducidas y/o brevemente definidas (lo que hemos denominado aquí un diccionario pasivo; véase el apartado 1.4.1) sino una obra compleja, con multitud de información útil para el usuario, que hace necesario entender dónde y cómo se origina, a quién va destinada y para qué ha sido confeccionada. Así pues, la etiqueta de “especializado” debe hacer al terminógrafo plantearse en primer lugar qué tipo de lenguaje va a tratar que justifique (o no) la adopción de tal denominación. Es por ello que cabe plantearse: ¿existe un límite claro y nítido entre lo general y lo especializado desde el punto de vista lingüístico? En este sentido existen tres posturas que Ahmad *et al.* (1995) resumen como:

- 1) La de aquellos que piensan que puede trazarse una línea clara y nítida entre lo general y lo específico/especializado en materia lingüística por considerar que los lenguajes especializados son códigos lingüísticos pero con unas unidades y reglas que los diferencian de la lengua general (Hoffmann 1998)
- 2) La de aquellos que consideran los lenguajes de especialidad como variantes del lenguaje general (Rondeau 1983; Rey 1979).
- 3) La de aquellos que consideran los lenguajes de especialidad como subconjuntos pragmáticos del lenguaje general y por lo tanto entienden que existe una confluencia de parte de las características del lenguaje general y otra parte de características propias. (Sager, Dungworth y McDonald 1980; Pitch y Draskau 1985).

Bergenholtz y Tarp (1995) también establecen una serie de posibles relaciones entre el lenguaje general y el especializado: por una parte, puede considerarse que, en mayor o menor medida, todos los lenguajes especializados usan el sistema del lenguaje general y pueden encontrarse dentro de él; por otra parte, desde una perspectiva léxica, puede entenderse que todas las expresiones del lenguaje general también se encuentran en los lenguajes de especialidad y por ello la lengua general es un subsistema de las lenguas de especialidad. Finalmente, desde una perspectiva comunicativa se puede entender que la lengua general y las lenguas de especialidad se usan en situaciones completamente distintas. De este modo los lenguajes de especialidad y la lengua común son fenómenos iguales y distintos a la vez.

Con todo, si hay algo que parece incuestionable es que dentro del ámbito del discurso especializado entendido en su globalidad, o dentro de los distintos discursos de especialidad existentes, se da una amplia gama de posibilidades que comprenden desde un discurso marcadamente especializado hasta una variedad de discurso fronterizo con la lengua común. Esto establece un continuum lingüístico entre “lo especializado” y lo “no-especializado” que parece indicar que la diferencia entre ambos es básicamente una cuestión de grado (mayor o menor especialización o abstracción). Existe además a ese respecto un claro indicador y es que, a medida que el lenguaje especializado se vuelve más preciso y conciso (especializado), puede ser entendido por un menor número de personas. Es por ello que para Cabré (1993: 148), aunque “algunos autores han concebido los lenguajes especializados por la temática como sistemas totalmente diferenciados de la lengua común [...] esta afirmación, en términos absolutos no parece muy adecuada” y no lo es porque este enfoque tan “individualista” sólo sería posible si, después de la comparación de dos textos, uno general y otro especializado, éstos no presentaran ningún tipo de coincidencia lingüística y por lo tanto la idea de que vienen de sistemas totalmente diferentes pudiera ser coherentemente defendida. Sin embargo, con toda seguridad, ambos textos mostrarán coincidencias lingüísticas muy significativas que no hacen sino confirmar y probar que han sido construidos en base a un sistema común, aún cuando los textos también presentarán diferencias importantes que podrán llevarnos a pensar que, en ciertos aspectos, provienen de sistemas distintos. Según Cabré (1993), lo que ocurre en este caso es que el código que

unifica ambos textos es el lenguaje común mientras que el código que los diversifica es el lenguaje especializado.

Consideramos pues la perspectiva a la vez integradora y diferenciadora de la tercera postura como la concepción de partida más plausible y coherente con la realidad lingüística existente, ya que ratifica lo afirmado por Cabré al establecer que el lenguaje común es el código unificador y que el lenguaje especializado es el código diversificador, algo que la autora resume en su conceptualización del lenguaje “marcado” y “no marcado” (Cabré 1993). Según esta dicotomía, el grupo de normas, unidades y restricciones que forman parte del conocimiento de la mayoría de los hablantes de una lengua constituye el llamado lenguaje común, cuyas unidades se usan en situaciones que podrían ser definidas como no marcadas. Por contra, los lenguajes especializados se usan para referirnos a un “conjunto de subcódigos (que coinciden parcialmente con el subcódigo del lenguaje común) caracterizados por una serie de particularidades especiales”. Estas características son específicas de cada lenguaje de especialidad y pueden referirse al tema, al tipo de interlocutor, la situación comunicativa, la intención del hablante, el medio en el que tiene lugar el intercambio comunicativo, etc. Por lo tanto, la situación en la que se dan los lenguajes de especialidad puede considerarse como marcada. Siguiendo a Gómez de Enterría (1998), podríamos decir que las lenguas de especialidad surgen dentro de la lengua común pero tienen como diferencia específica su objetivo último: cubrir las necesidades específicas de comunicación formal y funcional que se plantean en los diversos ámbitos técnicos y científicos.

Abundando en estas consideraciones, la plausibilidad de la postura 3 (véase apartado 1.1) como punto de partida de nuestra visión de la lengua general y los lenguajes de especialidad se sustenta también en el hecho de que el lenguaje especializado no puede ser concebido sin el lenguaje general pese a que este último sí puede sobrevivir por sí mismo. Es por ello que la importancia del lenguaje general para el lenguaje especializado es obvia como también lo es la influencia del lenguaje especializado en el lenguaje general, cada vez más notoria en la presencia de la ciencia y la tecnología en nuestro día a día (*e-mail*, *MP3/4*, *DVD*, *escanear*, *fax*, *antibiótico*..) y que hace que las lenguas de especialidad sean por ello (cada vez más) necesarias. En consecuencia, creemos que los lenguajes especializados son variedades funcionales

del lenguaje general –entendido como un sistema de signos simbólicos empleados en la intercomunicación–, parcialmente coincidentes con él y que se diferencian de éste por un grupo de características que se activan cuando confluyen una serie de circunstancias comunicativas, a saber: la especificidad o especialización del tema, el hecho de que los hablantes sean especialistas, la situación formal y profesional en que se dan y el hecho de que, desde un punto de vista discursivo, se esté tratando de transmitir algún tipo de información. Así pues, tal y como apunta Cabré (1990: 140) (véase la Figura 1), podríamos decir que los lenguajes de especialidad son variantes pragmáticas que están en relación de inclusión respecto a la lengua general y en relación de intersección respecto a la lengua común, con la que comparten rasgos y características y donde se produce un constante trasvase.

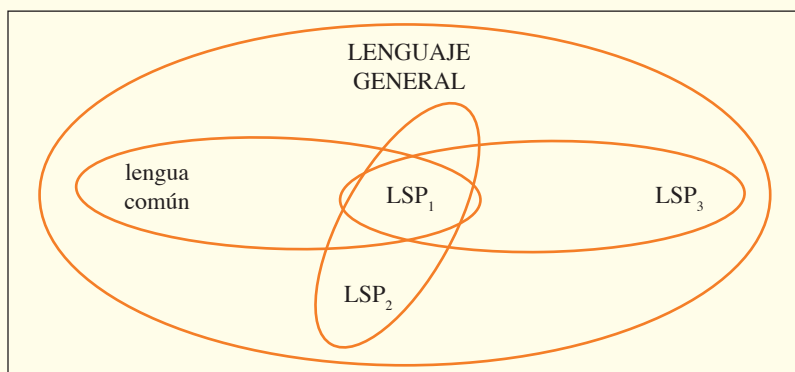


Figura 1. Relación entre lenguaje general, lengua común y lenguajes de especialidad (Cabré 1993: 140).

El sistema comunicativo que los interlocutores usan en un campo específico de conocimiento incluirá el lenguaje común (del que se usa la sintaxis, la morfología y parte del léxico y que podría definirse como el conjunto de recursos que usan la mayoría de hablantes de una lengua) y el subcódigo que es característico del campo de especialidad en sí y que contiene la terminología específica de dicho dominio, y cuya comprensión y correcto uso es la base/razón de ser de los diccionarios especializados. Así pues, los lenguajes de especialidad podrían definirse como subconjuntos de la lengua (variedades funcionales o registros), parcialmente coincidentes con la lengua común y que se utilizan como instrumento de comunicación formal y funcional entre especialistas de una determinada materia.

1.2 La terminología y la Teoría Comunicativa de la Terminología

Una vez entendida la relación entre los lenguajes especializados, el lenguaje general y la lengua común es necesario empezar a centrar nuestra atención en aspectos cada vez más concretos y para ello es fundamental entender el objeto último y específico de cualquier trabajo terminográfico. No obstante, esto no es posible sin comprender, en primer lugar, qué es la terminología. En general, el término ‘terminología’ suele referirse a tres conceptos diferentes: referido a la disciplina, es la base conceptual y el grupo de principios que gobiernan el estudio de los términos; en lo referente a la metodología comprende las directrices empleadas en el trabajo terminográfico (terminografía); en cuanto a la designación del grupo de términos de cada área temática específica, la terminología es el grupo de términos de un área de especialidad dada. El desarrollo de obras lexicográficas especializadas concierne a los tres aspectos mencionados y en base a ellos se ha estructurado también este artículo que tiene como objeto de estudio último la terminología entendida como vocabulario prototípico de un área de especialidad (3ª acepción).

La terminología siempre ha existido como el medio a través del cual la necesidad comunicativa de referirnos a cualquier concepto perteneciente a un campo de especialidad puede ser resuelta, y surge de la práctica y de la necesidad observada y experimentada por los científicos y técnicos de estandarizar o al menos “organizar” y “unificar” sus disciplinas desde un punto de vista conceptual y denominativo con el objetivo final de garantizar la comunicación profesional y una transmisión efectiva del conocimiento. Sin embargo, la terminología como disciplina tiene sus orígenes, aproximadamente, de 1930 a 1960, cuando aparecen los primeros textos teóricos de Wüster y Lotte, basados en la preparación de métodos de trabajo que consideraran el carácter sistemático de los términos. Según la Teoría General de la Terminología (TGT) de Wüster (1979), los conceptos existen independientemente de los términos que los designan y son, además, independientes del contexto. El radio de acción de la TGT no toma por tanto en consideración los aspectos pragmáticos y comunicativos inherentes al léxico de los distintos lenguajes de especialidad –y base en la concepción de dichos lenguajes adoptada aquí y expuesta en el apartado anterior– y afirma que los términos tienen una naturaleza biunívoca, habiendo una forma para cada concepto y un concepto

para cada forma. Esta teoría adopta un enfoque onomasiológico (del concepto al término) y tiene como objetivo principal la normalización. Presenta por ello esta teoría una serie de limitaciones importantes como son el reduccionismo en la manera de entender la terminología como disciplina y sus aplicaciones, la excesiva estandarización de la disciplina, el hecho de no concebir la realidad de manera realista desde un punto de vista comunicativo, el hecho de no ser capaz de dar cuenta de forma completa del fenómeno terminológico en la comunicación especializada y el hecho de confundir lo que es la terminología con lo que le gustaría que fuera (Cabré 1999), motivo por el cual se contempla como un punto de partida válido y reseñable para la evolución de la disciplina pero no como la teoría que la realidad lingüística nos aconseja tomar en consideración.

En la actualidad, sin embargo, desde una perspectiva social (aspecto claramente descuidado en la TGT), el carácter comunicativo de la terminología ha tratado de ser enfatizado a través de la importancia de su rol social como algo que trasciende su rol representacional o la estandarización de términos, que entronca con las necesidades reales derivadas de la terminología –básicamente, resolver necesidades sociales y optimizar la comunicación entre especialistas y profesionales– y que da lugar, según Cabré (1993), a una perspectiva pragmática y a una concepción aplicada de la terminología que encaja mejor con el espíritu de la sociedad actual.

La consideración de estos aspectos llevó con el tiempo al necesario desarrollo de teorías más acordes con la realidad terminológica y de los tiempos, como la Teoría Comunicativa de la Terminología (TCT) de Cabré (1999). Según esta teoría, clave en la propuesta lexicográfica especializada que aquí presentamos, los términos se analizan *in vivo*, caracterizados desde el “hábitat natural” (el uso real) en el que se dan en la comunicación especializada por lo que el texto especializado se convierte en el pilar básico en el que se fundamentan la teoría y la investigación terminológica adquiriendo éstas una dimensión textual que conduce a la adopción de una dimensión fraseológica. La TCT entiende la producción terminológica también desde un punto de vista descriptivo que concibe las unidades léxicas como ítems lingüísticos reales y no como entidades ideales y homogéneas. La terminología es abordada como un grupo de unidades denominativo-conceptuales que representan el conocimiento especializado dentro de un área

temática específica y se concibe como un vehículo para “hacer llegar” la comunicación profesional natural. La TCT, acorde a su rol de teoría social y comunicativa, asume la diversificación lingüística que depende del tipo de situación, del nivel de especialización de los textos, del emisor, del grado de formalidad, del objetivo final del discurso y del tema entre otros factores. Respecto al objeto de estudio de la terminología y más concretamente de la TCT, que luego resultará clave para entender los fundamentos subyacentes en la fase de extracción terminológica del método, partimos de la Hipótesis de la Unidad Léxica (Cabré 1999, 2000; Adelstein y Cabré 2000; Adelstein 2001; Cabré y Adelstein 2001), que plantea que la unidad léxica no es *per se* palabra ni término, sino una forma léxica que, de acuerdo con la situación comunicativa en la que se emplee, activa un valor especializado o no. Según esta hipótesis, las características y las situaciones discursivo-comunicativas son las que determinan la activación selectiva de información, por lo que, respecto a la posible dicotomía entre unidades terminológicas (o términos) y palabras, encontramos una disyuntiva similar a la planteada a la hora de discernir si los lenguajes de especialidad, la lengua común y la lengua general son sistemas distintos. Al igual que ocurría entonces y de acuerdo a la perspectiva adoptada en dicho nivel, consideraremos, como proclaman la TCT y la Hipótesis de la Unidad Léxica, que todas las unidades léxicas forman parte del lenguaje natural siendo, por tanto, objeto de estudio de la lingüística, y que algunas activan su carácter especializado en contextos específicos. Así pues, los términos son unidades del lenguaje natural que pueden activar conocimiento especializado dependiendo de la situación comunicativa por lo que forman parte, al igual que las unidades léxicas más generales, del léxico y por tanto de la lingüística y del lenguaje natural por lo que una distinción abrupta entre ambas no responde a un enfoque integrador y comunicativo como el que aquí pretende mostrarse.

Es por todo lo dicho hasta el momento que el uso de corpus lingüísticos se hace fundamental en el desarrollo de diccionarios especializados activos de enfoque comunicativo según la lexicografía especializada, que permitirán el estudio de una terminología esencialmente representacional o comunicativa, esto es, de una terminología “*in vivo*” (natural, espontánea) frente a concepciones representacionales que estudian la terminología “*in vitro*” (estandarizada) siguiendo los

principios de univocidad y monosemia postulados por la teoría clásica. En consecuencia, la TCT y por tanto la propuesta aquí presentada aboga por observar los términos tal y como aparecen en las diferentes situaciones comunicativas y por ende sujetos a cambio, no pudiendo por tanto concebir otro modo de aproximarse y estudiar el término que no sea a través de corpus lingüísticos y de enfoques semasiológicos que van del término (según se da en el corpus de estudio) al concepto.

1.3 La lingüística de corpus: una aproximación empírica al análisis y la descripción de la lengua en uso

Parece incuestionable que toda ciencia que tenga como objeto la formulación, rechazo o corroboración de hipótesis para dar cuenta de ciertos aspectos de la realidad deba estar fuertemente basada en los datos ofrecidos por esa misma realidad. Este tipo de datos se conocen como datos empíricos y las ciencias humanas también pueden beneficiarse de ellos (Edo Marzá 2009).

Los datos empíricos proceden de y reflejan el aspecto de la realidad que tratan de describir, haciendo posible la validación o rechazo de hipótesis. Así pues, la lingüística necesita elementos de “referencia real y verdadera” tanto como las ciencias naturales, con el fin de proporcionar una visión real y realista del lenguaje así como para confirmar o refutar hipótesis sobre el mismo, sobre su uso y sobre su comportamiento en contexto. Esta visión y aproximación real sólo puede darse de manera efectiva mediante el uso de corpus textuales. De modo breve, podría decirse que, tal y como plantea Sánchez (1995), un corpus lingüístico es un conjunto de datos lingüísticos pertenecientes al uso oral o escrito de la lengua, o ambos, que han sido sistematizados según determinados criterios suficientemente extensos en amplitud y profundidad de manera que sean representativos del total del uso lingüístico o de algunos de sus ámbitos, y dispuestos de tal modo que puedan ser procesados mediante ordenador con el fin de obtener resultados varios y útiles para la descripción y el análisis. Hoy en día, tal y como ya apuntaba Sánchez, esta última cualidad que podríamos denominar como el hecho de ser “electrónico” o procesable parece necesaria e incluso inherente al concepto mismo de corpus.

La lingüística de corpus se vio fuertemente cuestionada con la llegada de Noam Chomsky, principal representante del Generativismo y que cambió la dirección empírica que estaba adoptando la lingüística

(con la temprana lingüística de corpus) hacia lo que se conoce como “racionalismo”, un paradigma según el cual el lenguaje debía ser observado a través del juicio intuitivo e introspectivo del investigador. De acuerdo con esto, un corpus nunca podría ser una herramienta útil para los lingüistas porque estos deben buscar un modelo de “competencia del lenguaje” (*language competence*) más que de “uso del lenguaje” (*language performance*). Además, como el número de enunciados de un lenguaje es infinito, la existencia de un repertorio finito de datos lo suficientemente adecuado para explicar todos los mecanismos de producción lingüística no es posible, lo que convertía a los corpus en inservibles según este paradigma. Sin embargo, pese a las insuficiencias e inconsistencias mostradas por la temprana lingüística de corpus y que originaron su cuestionamiento por parte de autores como Chomsky, la lingüística de corpus nunca fue totalmente abandonada. De forma análoga a lo ocurrido con la terminología, a finales de los años 70 y comienzos de los 80, los ordenadores personales se convirtieron en herramientas de fácil acceso y fue entonces cuando el uso de corpus textuales (y de las herramientas necesarias para su explotación y análisis) vivió una auténtica eclosión que sigue en auge hoy en día. Tal y como apuntan Rafel y Soler (2003), la introducción de los medios informáticos en la constitución, el procesamiento y la explotación de corpus textuales ha hecho posible el tratamiento de grandes volúmenes de información de una manera sistemática y objetiva, sin la intervención de factores externos que no deben formar parte de la investigación, como por ejemplo la intuición del lingüista. De hecho, la informática ofrece innumerables recursos y herramientas clave en lexicografía especializada y que hacen fáciles, rápidas e incluso precisas las tareas más repetitivas de las etapas del trabajo terminográfico, tal y como veremos en los apartados siguientes.

A la hora de diseñar un corpus de trabajo *ad hoc* para la extracción y el análisis de la terminología, es importante tener en cuenta cuatro factores principales: el tamaño, la representatividad (marcada por la diversidad y el equilibrio), la procesabilidad o explotabilidad y la fiabilidad de los datos. Por considerar los dos últimos parámetros prácticamente autodefinitorios y debido a las limitaciones en cuanto a espacio de este estudio, analizamos en más detalle aquí únicamente las dos primeras.

En lo referente al tamaño, el hecho de tener un número total de textos muy bajo (aunque un texto sea muy largo e incluya una gran cantidad

de palabras) puede tener una influencia negativa en los resultados de un análisis. Se corre el riesgo de que las características lingüísticas personales de un autor concreto o la preponderancia textual de una subárea dentro de un corpus se tomen como 100% representativas y características del lenguaje o registro analizado cuando, en realidad, simplemente reflejarán una falta de equilibrio en el diseño del corpus y por tanto no constituirán rasgos generalizables. En consecuencia, de acuerdo con Biber (1990), en un corpus deben incluirse textos suficientes en cada categoría para abarcar la variación que se da entre hablantes o autores. Siguiendo por ejemplo a Sinclair (1996), diremos que suele considerarse necesario tener acceso a corpus textuales grandes porque los ejemplos buenos y representativos son difíciles de encontrar. Es por todo ello que el terminógrafo necesita tener un número elevado de ocurrencias disponibles para ser analizadas si quiere estudiar con ciertas garantías de representatividad y equilibrio el comportamiento de las palabras en el texto.

En lo referente a la representatividad, el terminógrafo debe tener los pies en la tierra y ser consciente de que, tal y como señalan Biber, Conrad y Reppen (1998) es del todo imposible conocer la extensión o el alcance total de la variedad en el lenguaje o todas las variables contextuales que necesitan ser cubiertas para verdaderamente tener un corpus 100% representativo. Es por ello que, según estos mismos autores, el diseño adecuado de un corpus depende de lo que dicho corpus pretenda representar. "Representatividad", de acuerdo con Tognini-Bonelli (2001), significa representativo de una cierta población y que las conclusiones derivadas del análisis del corpus serán ampliamente aplicables a una muestra de población mayor entendida como un todo. Es por ello que un corpus representativo de un dominio debe incluir, en general, el mismo número de muestras de 4 tipos distintos de textos que dan cuenta de o representan los distintos niveles de especialización (Gómez y Vargas 2004): muy especializados, especializados, didácticos y divulgativos. Al mismo tiempo esta clasificación está estrechamente relacionada con el nivel de especialización del autor y de los destinatarios o receptores (que pueden ser, en general, especialistas, semi-especialistas, estudiantes y público en general) (Gómez y Vargas 2004).

Con todo, con un enfoque basado en corpus para el estudio del lenguaje, el terminógrafo ha de ser consciente de que es más que probable

que algunas ocurrencias queden excluidas por ser “raras” y sin embargo, otras que deberían aparecer por ser mucho más comunes también queden excluidas por meras cuestiones de azar. Del mismo modo, sin embargo, también puede darse el caso de que ciertas ocurrencias extremadamente raras aparezcan varias veces en un corpus, dando como resultado una frecuencia que no es exactamente representativa de, por ejemplo, una variedad del lenguaje. Pese a ello, una aproximación a la lingüística de corpus y al uso de corpus textuales resulta indispensable en la propuesta lexicográfica especializada aquí presentada y que se completa, desde el punto de vista teórico, con el siguiente apartado.

1.4 La lexicografía especializada: un paso más allá de la terminografía

A lo largo de los subapartados anteriores hemos tratado de presentar de forma los fundamentos teóricos básicos necesarios para comprender y llevar a cabo de manera efectiva la elaboración de diccionarios especializados desde un enfoque lexicográfico especializado. No obstante, es cierto que en varios puntos de este estudio se han utilizado los términos terminografía o terminográfico como sinónimos de lexicografía especializada o lexicográfico especializado, por lo que es necesario en este punto plantearse la siguiente cuestión: ¿es realmente la terminografía lexicografía especializada? ¿En qué consiste el enfoque lexicográfico especializado?

Según lo planteado aquí –desde un punto de vista tanto teórico como práctico o metodológico– y que consideramos el más adecuado para la elaboración de diccionarios especializados, la visión de que la terminografía no es lexicografía especializada, basada en una concepción de la terminología como una disciplina independiente y diferenciada de la lexicología y con un objeto de estudio particular, no tendría sentido. Para aquellos defensores de esta opción, ambas disciplinas no se relacionarían ni a nivel teórico ni práctico o metodológico, pero esto no es lo que se desprende del enfoque aquí propuesto. Sin embargo, la pregunta se respondería positivamente si, como en este caso, se considera que la terminología puede explicarse desde la lingüística y, más exactamente desde la lexicología y que tanto el objeto de estudio –en caso de querer parcelarlo– de la lexicología como de la terminología pertenecen al lenguaje natural y por tanto entran dentro del campo de estudio de la lingüística; además, la condición

de término puede explicarse a partir de la activación de una serie de rasgos semánticos de las unidades léxicas cuando éstas se utilizan en contextos de especialidad y la distinción en la metodología utilizada en la elaboración de los respectivos productos terminográficos y lexicográficos (la tradicional dicotomía entre los enfoques onomasiológico y semasiológico respectivamente) tampoco es válida según la perspectiva adoptada aquí y que hace uso de corpus textuales. Es por ello que la tradicional distinción entre lexicología y terminología y lexicografía y terminografía no resulta válida aquí por ser distinciones artificiales e innecesarias. En palabras de Santamaría (2006), desde la actual concepción de la terminología de base lingüístico-comunicativa –la misma que venimos defendiendo a lo largo de este estudio– se aboga por un acercamiento a la terminografía desde la lexicografía y se apuesta por una denominación común para aglutinar en un único marco teórico diferencias que no son reales entre ambas disciplinas: *lexicografía especializada*, esto es, una lexicografía que estudia las unidades temáticamente especializadas y que abarca todo tipo de diccionarios o inventarios léxicos especializados por la temática.

La lexicografía especializada aboga por una visión de la lengua como continuum que tampoco permite la clasificación estanca e independiente de disciplinas que son necesariamente afines. Si entendiéramos la lexicología (y su vertiente aplicada, la lexicografía) y la terminología (y su vertiente aplicada la terminografía) como disciplinas lingüísticas independientes y entendiéramos que, tal y como afirmaba Wüster (1979), ambas tratan objetos de estudio distintos –las palabras en el caso de la lexicología y los conceptos (representados por términos) en el caso de la terminología– y utilizan enfoques distintos –la terminología debe ir del concepto al término (onomasiológico) y la lexicología del término al concepto (semasiológico)– entonces tendría sentido separar taxativamente la lexicología de la terminología (y lo mismo respecto a sus vertientes aplicadas). No obstante, el estudio de las lenguas de especialidad, el análisis de las unidades léxicas en contexto y en uso, el fin comunicativo del lenguaje y, por qué no, nuestro propio conocimiento del mundo, nos indican que eso no es así. Es por ello que desde una perspectiva lingüística y comunicativa puede afirmarse que la terminología es lexicología especializada o una lexicología de las unidades léxicas especializadas por pertenecer al léxico del hablante y por tanto también a la lingüística y al lenguaje

natural. De ahí que consideremos la etiqueta “lexicografía especializada” como más exacta, integradora y comunicativa en su enfoque que la etiqueta “terminografía”, tal y como ya hicieran otros autores como Bergenholtz y Tarp (1995) y Lerat (1995) que ya apostaban por una denominación común para aglutinar en un único marco teórico diferencias que no son reales, entre lexicología y terminología: la lexicografía especializada o la lexicografía terminológica, esto es, una lexicografía especializada basada en los conceptos de lenguas de especialidad y lengua general (Rey 1995: 119).

Así pues, abogamos por una visión de la lexicografía especializada como un cuasi-sinónimo de la terminografía pero que trasciende los principios y concepciones terminográficos clásicos para adaptarse a enfoques más realistas en el tratamiento de la terminología. Es cierto que los términos, o lo que Cabré (1993) acertadamente denomina Unidades de Conocimiento Especializado (UCE), son el primer objeto de estudio de la terminología como disciplina, pero también lo es que los “términos” no forman parte de un sistema independiente de las palabras, sino que forman parte del léxico del hablante y se distinguen de las unidades léxicas no especializadas por su contenido preciso y bien delimitado en el contexto de uso de una materia especializada. Es por ello que la terminología suele entenderse como el “mero” y tradicional estudio de los términos (posible gracias a la terminografía) que, a su vez, puede ser complementado por la mayor proyección y visión aunadora de la lexicografía especializada para la elaboración de diccionarios especializados, orientados al usuario, con un enfoque más comunicativo y social y de fácil manejo y comprensión (*user-friendly*). Como resultado, el enfoque lexicográfico especializado que aquí presentamos se completa con el concepto de diccionario activo –que tomará forma, junto con el resto de principios teóricos expuestos– en la metodología para la elaboración de diccionarios especializados y sus correspondientes fundamentos prácticos expuestos en el apartado 2.

1.4.1 Los diccionarios activos en lexicografía especializada

Los diccionarios especializados activos se originan en una observación muy simple: hace algunos años, los terminógrafos y los lingüistas empezaron a revisar diccionarios especializados ya elaborados y detectaron una falta de entradas con categorías distintas al nombre, así

como una falta de contextos e informaciones que ayudaran a ilustrar el uso de los términos.

Hoy en día, se espera una mayor atención a las palabras como elementos oracionales y textuales que como elementos (artificialmente) aislados. Es por ello que, de acuerdo a los principios básicos de la TCT, de la lingüística de corpus y de los diccionarios activos, y por tanto dentro del enfoque especializado lexicográfico, los diccionarios deben diseñarse para cumplir las necesidades específicas de los grupos de usuarios para los que están diseñados, por lo que existe la necesidad de proveer a la terminología de materiales que satisfagan las necesidades comunicativas que surgen en las áreas de especialidad y que sean el resultado de estudiar los términos *in vivo*.

La idea de lo que un diccionario activo implica en comparación con los clásicos diccionarios que podríamos denominar “pasivos” ha sido recogida en la Figura 2, adaptada de Posteguillo y Piqué (2005), y en la que podemos observar la importancia de proporcionar al usuario la información necesaria tanto a nivel receptivo como productivo y traductológico para entender los términos en todo su espectro, para lo cual se propone incluir en las entradas lexicográficas especializadas información técnica, gramatical, colocacional, contextual y relacional.

Diccionario especializado pasivo	Diccionario especializado activo
TÉRMINO EN LENGUA DE ORIGEN Traducción a la lengua meta (a veces con información técnica)	TÉRMINO EN LENGUA DE ORIGEN Traducción a la lengua meta con: Información gramatical Información técnica Colocados Contextos/ejemplos de uso Área temática Referencias cruzadas / sinónimos
TÉRMINO EN LENGUA META Traducción a la lengua origen	TÉRMINO EN LENGUA META Traducción a la lengua origen con: Información gramatical Información técnica Colocados Contextos/ejemplos de uso Área temática Referencias cruzadas / sinónimos

Figura 2. Diccionarios pasivos vs. Diccionarios activos
(adaptada de Posteguillo y Piqué (2005)).

Tal y como apunta Tarp (2005), la función de un diccionario en el sentido más amplio del término es suministrar los datos lexicográficos

de los cuales un tipo específico de usuario que se encuentra en un tipo específico de situación extra-lexicográfica de uso pueda obtener la información necesaria para satisfacer el tipo específico de necesidades que puedan surgir en este tipo de situación. En el caso de los diccionarios especializados activos, creemos que el enfoque aquí propuesto puede y debe proporcionar al usuario final la información necesaria a la hora de afrontar cualquier situación comunicativa derivada de una necesidad terminológica.

Para concluir, diremos que toda práctica terminológica presupone, según Cabré (1999), el dominio de cuatro grandes competencias: la competencia cognitiva, correspondiente al conocimiento del ámbito especializado que va a ser objeto de estudio; la competencia lingüística, que abarca el conocimiento de la lengua o las lenguas sobre las que se quiere realizar el trabajo; la competencia metodológica o el conocimiento del método para la elaboración de productos terminológicos y la competencia sociofuncional, referida a las características que debe tener el trabajo terminológico para lograr los fines que persigue y ser adecuado a los destinatarios a los que se dirige. Es precisamente esta competencia sociofuncional la que determina que, en palabras de Edo Marzá (2009), la dimensión lingüística de todo trabajo lexicográfico especializado deba entenderse siempre como aplicable e insertada en el mundo real y en un contexto socio-económico y profesional concreto del que forma parte el usuario de la obra y en el que se originan sus necesidades. Del mismo modo, la ya citada competencia metodológica, que trataremos de forma más detallada en el punto 2, le debe gran parte de su potencial y desarrollo actual a otro factor clave en la lexicografía especializada: la terminótica. Sin ella, hablar de lexicografía especializada sería hoy en día impensable porque las herramientas terminóticas están presentes en prácticamente todos los estadios de elaboración de los diccionarios especializados. El procesamiento de grandes corpus textuales, la explotación y análisis de esos corpus, así como el almacenamiento y gestión de los datos a través de bases de datos terminológicas serían hoy en día inconcebibles a la escala que conocemos de no ser por la terminótica.

Por último, apuntar que, si bien no se ha podido entrar a valorar en detalle muchos de los aspectos que conforman esta propuesta, sí se ha tratado de exponer de forma clara las bases teóricas que debieran sustentar, desde nuestro punto de vista, la elaboración de cualquier

diccionario especializado desde la perspectiva lexicográfica especializada. Una vez vista esta fundamentación teórica, se hace pues necesaria la exposición de la fundamentación práctica que conlleva necesariamente el planteamiento de una metodología de trabajo conducente al fin último de la lexicografía especializada: la elaboración de diccionarios especializados de calidad y utilidad.

2. FUNDAMENTOS METODOLÓGICOS PARA EL DESARROLLO DE DICCIONARIOS ESPECIALIZADOS

En cuanto a los fundamentos prácticos que aquí proponemos para el desarrollo de diccionarios desde la lexicografía especializada, cabe señalar la organización de estos en base a una propuesta de método compuesta de ocho etapas básicas que, entendemos, deben guiar en general la elaboración de diccionarios de especialidad de calidad según la perspectiva teórica expuesta en el apartado 1. Estas etapas, basadas en la metodología propuesta por Auger y Rousseau (1987) y en la adaptación a las nuevas tecnologías hecha de la misma por Gómez y Vargas (2004), son: (1) definición del trabajo; (2) elaboración del árbol de campo; (3) preparación y compilación del corpus y de otros recursos documentales; (4) gestión documental del corpus; (5) explotación del corpus: extracción y análisis terminológico; (6) procesamiento y almacenamiento de los datos; (7) revisión y normalización y (8) edición.

Para ilustrar las diversas fases propuestas en el método hemos utilizado el *Diccionario Bilingüe (Inglés-Español, Español-Inglés) de la Industria Cerámica*, un proyecto interuniversitario del que la autora de este artículo es también coautora junto a la Dra. Chelo Vargas (Universidad de Alicante, España) y al Dr. Santiago Posteguillo (Universitat Jaume I (Castellón) España) y que, pese a no estar todavía acabado, puede ejemplificar desde una obra real los aspectos tratados.

2.1 Definición del trabajo

Es importante partir siempre de la base de que cualquier especialista o lingüista que esté a punto de empezar un trabajo terminológico debe tener un cierto conocimiento sobre 4 aspectos clave del tema de su trabajo: los contenidos del tema/materia, la documentación disponible, el contexto profesional en el que se desarrolla el tema y la situación sociolingüística del dominio.

Además, tal y como apuntan Cabré y Estopà (2006), los trabajos terminológicos no se pueden realizar en abstracto sino que deben siempre estar motivados por una necesidad profesional; deben respetar el supuesto de adecuación cubriendo tanto las necesidades profesionales como las funciones para las que se utilizarán (lo que influirá en cada fase del trabajo), y resultarán adecuados y coherentes siempre respecto a una situación comunicativa, esto es, a una actividad profesional, un ámbito temático, un contexto sociolingüístico, una función lingüística y unos usuarios prototípicos.

Es por ello que la elaboración de cualquier diccionario debe partir siempre de una propuesta ordenada, organizada, lógica, coherente, preparada y consensuada por el equipo si lo hubiere. En consecuencia, se propone en esta etapa inicial la redacción de un documento interno de trabajo en el que se establecerán claramente los principales aspectos que definen el tipo de trabajo terminográfico a realizar, a saber, los destinatarios o usuarios –que determinarán la manera de presentar la información–, las lenguas de trabajo –pudiendo ser las obras monolingües, bilingües o plurilingües–, los objetivos o finalidad del diccionario así como el equipo de trabajo y los asesores, siendo éstos los encargados de garantizar la calidad lingüística y técnica de la obra. El equipo de trabajo deberá necesariamente contar con o estar formado, pues, por miembros con conocimientos, además de lexicográficos, de índole lingüística más general, conocimientos de la ciencia o materia tratada, conocimientos informáticos y terminológicos e, idóneamente, conocimientos de maquetación, si bien hoy en día muchos programas generan un maquetado prácticamente automático de las entradas. Se determinarán además en esta fase inicial los recursos disponibles en cuanto a equipamiento informático (PC, escáner, fotocopidora, acceso a Internet...) y programas (por ejemplo, *OmniPage Pro*, *WordSmith Tools*, *TermStar*, *GenDic*, etc.). Además, es crucial establecer una correcta delimitación de las funciones lingüísticas, lo que conllevará una necesaria fijación de las funciones del futuro diccionario como descriptivas (fin referencial y centrado en los términos que usan los especialistas en su comunicación profesional) o prescriptivas (fin prescriptivo centrado en orientar el uso que los especialistas deben hacer de los términos).

Por último, en esta primera fase de definición del trabajo se establecerá una delimitación temática inicial del área de especialidad a

tratar; esto implica en primer lugar y de acuerdo a las características intrínsecas del trabajo, decidir si se va a realizar un trabajo sistemático –que abarca un área o subárea de especialidad– o puntual –centrado en un solo término o un pequeño conjunto de los mismos– para posteriormente establecer de manera preliminar la dimensión y el alcance conceptual y terminológico del trabajo y la exclusión de los aspectos que, a priori, no sean relevantes para el trabajo a realizar, constituyendo una primera aproximación al árbol de campo que veremos en el apartado 2.2. La Tabla 3 muestra la delimitación temática preliminar establecida para el diccionario de la industria cerámica.

Caract. Materias primas	Procesos Productivos	Caract. del producto acabado	Aplicaciones	Otros
Materias primas para la industria cerámica	Extracción de las materias primas	Ensayos	Aplicaciones interiores	Organismos e instituciones
Propiedades de las materias primas	Transformación de las materias primas	Comercialización	Aplicaciones exteriores	General
	Procesos productivos para la obtención del producto		Aplicaciones decorativas	
	Medidas de seguridad		Mantenimiento	

Tabla 3. Delimitación temática preliminar del dominio cerámico industrial.

2.2 Elaboración del árbol/diagrama de campo

En base a la definición del trabajo establecida y a su correspondiente delimitación temática preliminar, el lexicógrafo especializado debe saber reconocer, delimitar y estructurar dentro de límites razonables el dominio de especialidad a tratar. Es por ello que se elaborará una representación gráfica definitiva del campo de estudio que recogerá la organización conceptual del área. A esta representación gráfica la llamaremos “árbol de campo”. Un árbol o diagrama de campo es una estructura organizada y organizativa de clases nocionales o áreas temáticas que conforman un dominio de especialidad y que mantienen entre sí una serie de relaciones basadas en las características que

comparten. Este tipo de diagrama es una representación gráfica de los campos y subcampos que están comprendidos en una determinada actividad/materia. Es una herramienta útil debido a su naturaleza asociativa, relacional y clasificatoria. Las relaciones que muestra hacen del árbol de campo un todo coherente determinado por la noción de campo semántico que, de acuerdo con Auger y Rousseau (1987), es un grupo de nociones que pueden ser agrupadas bajo el encabezamiento de una noción clave.

El árbol de campo es, pues, una herramienta para establecer un sistema ontológico, es decir, una lista estructurada dentro de la cual las diferentes actividades pertenecientes al sector analizado puedan ser incluidas y clasificadas y la estructura mental del lexicógrafo pueda organizarse de manera significativa. Las funciones de este árbol de campo son pues, básicamente, delimitar el ámbito de trabajo especificando sus partes y su relación con el todo; mostrar las relaciones jerárquicas establecidas entre las unidades terminológicas; elaborar y determinar una lista de categorías nocionales que se utilizarán como guía para seleccionar los términos durante la extracción terminológica (ver apartado 2.5) y ubicar los términos dentro de un grupo estructurado que ayude en la correcta clasificación y tratamiento de las unidades terminológicas.

Una estructuración conceptual bien elaborada debe tener, según Cabré (1993), las siguientes características: debe cubrir/comprender todo el campo de trabajo; debe contener o incluir todas las ramas en que se estructura el sector analizado; debe presentar o representar la relación existente entre las distintas ramas y dentro de ellas; no debe incluir bloques conceptuales no pertinentes; debe estar estructurada de tal manera que el mismo concepto no se repita en ramas distintas, y no debe incluir bloques conceptuales no productivos, es decir, aquellos que no conduzcan a un grupo de conceptos (y a sus respectivas denominaciones). Además, el lexicógrafo debe tener siempre en cuenta que una estructuración de este tipo debe ser equilibrada, detallada y completa en relación al espacio conceptual que pretende cubrir el diccionario, por lo que puede servirse de tesauros especializados y posibles clasificaciones temáticas previas. Para ilustrar lo que es un árbol de campo, la figura 4 muestra el árbol de campo en español (recordemos que es un diccionario bilingüe) de la industria cerámica y azulejera.

1. Caracterización	1.1 Materias primas [MATPRI] 1.2 Propiedades de las materias primas [PROPMATPRI] 1.3 Geología/Mineralogía/Cristalografía [GEO-MIN-CRIS] 1.4 Química [QUIM]	
2. Procesos productivos	2.1 Extracción de las materias primas	2.1.1 Explotación y cantería [EXPL-CANT] 2.1.2 Maquinaria de extracción y explotación [MAQEXTRAC]
	2.2 Transformación de las materias primas	2.2.1 Procesos de transformación de las materias primas [TRANSPROC] 2.2.2 Materiales para la transformación [TRANSMTRLS] 2.2.3 Maquinaria de transformación [MAQ-TRANS]
	2.3 Transformación para la obtención del producto	2.3.1 Procesos productivos para la elaboración del producto [ELABPROC] 2.3.2 Esmaltes y productos relacionados [ESMREL] 2.3.3 Materiales [MTRLS] 2.3.4 Maquinaria de elaboración [MAQELAB] 2.3.5 Instrumentos de elaboración [INSTR] 2.3.6 Productos resultantes/finales [PRODFIN] 2.3.7 Subproductos [SUBPROD] 2.3.8 Ensayos [ENSAYO] 2.3.9 Propiedades químicas y físicas [PROPQUIM-FIS] 2.3.10 Calidad [CALIDAD] 2.3.11 Diseño [DISEÑO]
	2.4 Comercialización	2.4.1 Gestión [GESTIÓN] 2.4.2 Comercio [COM] 2.4.3 Distribución [DISTRIB]
	2.5 Medidas de seguridad y salud laboral [SEG]	
	2.6 Gestión medioambiental [MEDAMB]	
3. Aplicaciones	3.1 Mantenimiento [MANTNMO] 3.2 Instalación [INSTALA] 3.3 Construcción [CONST]	
4. Temas relacionados	4.1 Industrias afines [INDAFIN] 4.2 Medidas [MEDIDA] 4.3 Organismos e instituciones [ORG-INST]	

Figura 4. Árbol de campo en español elaborado para el diccionario de la industria cerámica.

Se recomienda la inclusión de esta estructura organizativa en el propio diccionario para facilitar al usuario la comprensión del alcance conceptual del mismo así como la perspectiva adoptada con respecto a la organización de la materia. Los códigos que aparecen dentro de

cada área y subárea de dicha estructura son las codificaciones establecidas para referirse a los distintos campos semánticos y constituyen, a su vez, uno de los campos fundamentales de las fichas terminológicas diseñadas, siendo uno de los primeros elementos dentro de la microestructura de las entradas y subentradas del diccionario (véase el apartado 3).

Por último es conveniente apuntar que este tipo de organización conceptual también puede hacerse a través de editores de mapas gratuitos tipo CmapTools (<http://cmap.ihmc.us/>), DigiDocMap (<http://www.mapasconceptuales.com>), o FreeMind (http://freemind.sourceforge.net/wiki/index.php/Main_Page), entre otros, pero, en cualquier caso, lo importante en esta fase del trabajo es que el lexicógrafo logre adquirir y reflejar una estructura conceptual clara del dominio o subdominio a tratar, sea por medios más o menos tradicionales.

2.3 Preparación y compilación del corpus y de otros recursos documentales

Tras la familiarización con el dominio fruto de las dos primeras etapas, la tercera gran fase aquí propuesta será la de preparación y compilación del corpus de trabajo y de otros recursos documentales tanto de consulta como de ayuda o soporte al trabajo. Estos recursos pueden ser terminográficos (diccionarios, enciclopedias, terminologías, bases de datos, tesauros, clasificaciones, etc.) o puede tratarse de documentos con información sobre la especialidad que en vez de ser incluidos en el corpus son utilizados como documentos de consulta (tesis doctorales, actas de congresos, etc.)

En el apartado 1.3 de este artículo ya se ha hecho referencia a la importancia de los corpus textuales para la descripción de la lengua en uso así como a la importancia de compilar corpus representativos, equilibrados y de un tamaño adecuado para tales fines. Además, entre los criterios para la compilación de un buen corpus de trabajo, Codina (2000) añade el considerar la fiabilidad de los documentos electrónicos atendiendo a criterios de rigor, sistematicidad, autoría, actualidad y originalidad. Abundando un poco más en este aspecto, aconsejamos la evaluación de los documentos a incluir en el corpus (no sólo los electrónicos) en base a su fecha de publicación, a la actualidad de la bibliografía que presentan, a su grado de pertenencia al dominio de especialidad, a la reputación de su autor, al grado de sistematización

de la información, a la calidad lingüística del texto, al idioma original de la obra y a la naturaleza del texto. Si bien el corpus perfecto no existe, sí es cierto que un corpus bien compilado es fundamental para la obtención de una obra lexicográfica de calidad por constituir la “materia prima” básica de ese tipo de obras.

Simplemente nos quedaría apuntar que para la compilación de un corpus de textos especializados fiable, representativo y equilibrado, se propone incluir, en general, textos de distintos niveles de especialización y naturaleza (o géneros representativos del dominio) como fueron, en el caso del diccionario de la industria cerámica, monografías, publicaciones periódicas, artículos de revista, tesis, informes, textos legales o normas (ISO, AENOR, etc.), manuales, actas de congresos y coloquios, publicaciones de divulgación, prospectos y folletos publicitarios, etc.

2.4 Gestión documental del corpus

Una vez el corpus ha sido compilado en base a los criterios de calidad y representatividad establecidos, es necesario convertirlo en procesable y compatible con los programas a utilizar en las etapas siguientes. Es por ello que habrá que almacenar las muestras como una colección de archivos procesables por ordenador y compatibles con el programa elegido para la fase siguiente, la de explotación del corpus (para la que, en nuestro caso, utilizamos el programa WordSmith Tools). En consecuencia, los textos en formato electrónico fueron automáticamente incorporados al ordenador en formato *txt* mientras que los textos impresos debieron ser digitalizados por medio de un proceso de escaneado y OCR. En el caso concreto del diccionario de la industria cerámica, el programa de OCR utilizado a tales efectos fue Omnipage, con el cual, tras el escaneado, se seleccionaron manualmente las secciones de texto a las que queríamos o nos interesaba aplicar el proceso de OCR y mediante éste el programa reconoció ópticamente los caracteres de las muestras textuales, pudiendo guardar los textos también como archivos en extensión *.txt*. Una vez realizada la gestión documental del corpus, las muestras textuales están listas para trabajar con ellas en posteriores etapas porque ya hemos compilado un corpus totalmente electrónico y compatible con el programa a utilizar en la fase siguiente.

2.5 Explotación del corpus: extracción y análisis terminológico

Una vez que el corpus está listo para ser explotado y que los aspectos previos del trabajo han quedado bien definidos y establecidos, el siguiente paso consistirá en elegir un programa de explotación de corpus textuales (o programa o *software* de concordancias) para poder efectuar, en primera instancia, una extracción terminológica preliminar. Son muchos los programas de este tipo disponibles en el mercado, tanto en forma de software libre como de programas comerciales pero destacaremos aquí la posibilidad de usar MonoConc, AntConc o WordSmith Tools (WST) en cualquiera de sus versiones si bien, obviamente, cuanto más actualizadas, más perfeccionadas. MonoConc y AntConc son especialmente atractivos por poseer una interfaz de usuario fácil y cómoda (*user-friendly*), y por ser capaces de realizar con total eficacia las tareas consideradas fundamentales en cualquier software de concordancias (generar listas de frecuencia, analizar concordancias, co-ocurrencias, etc.) Sin embargo, ni MonoConc ni AntConc tienen el potencial de un programa como WST ni pueden procesar cantidades tan grandes de texto. WST, por su parte, es menos *user-friendly* a primera vista que otros programas pero mucho más indicado para el lexicógrafo experto que busca optimizar su tiempo, personalizar al máximo los entornos y opciones de trabajo y sacarle el mayor provecho a su investigación.

Para esta explotación del corpus de trabajo, el enfoque lexicográfico especializado propuesto aquí hace lógica la combinación y el uso de varios tipos de resultados: resultados a nivel de frecuencia y grado de ligazón (información mutua) entre términos, resultados a nivel de saliencia (*keyness*), y resultados a nivel textual, contextual y co-textual. De acuerdo con esto, en la fase de extracción y análisis terminológicos, los datos correspondientes a la generación y análisis de los datos de frecuencia y de información mutua se obtienen con la herramienta de WST llamada Wordlist. El segundo aspecto, el de saliencia, se analiza con los resultados obtenidos con la herramienta Keywords, mientras que el análisis textual, contextual y co-textual lo realizan básicamente las distintas aplicaciones de la herramienta Concord. La Tabla 5 presenta de forma resumida todo el proceso de extracción y análisis terminológico aquí propuesto tal y como se realizó para el diccionario de la industria cerámica y llevado a cabo con WordSmith Tools en su versión 5.0. En los párrafos siguientes

explicamos esos pasos con mayor detalle, si bien de manera algo sucinta debido a las limitaciones de espacio lógicas en cualquier artículo.

Resumen del proceso de extracción y análisis terminológico con WordSmith Tools		
WordList (1ª parte)	Preselección (fase 1)	Comparación y combinación de resultados de frecuencia y saliencia con el fin de obtener una primera lista de términos monoléxicos potenciales.
	Generación de listados de frecuencia monoléxicos (fase 2)	
KeyWords	Generación de listas de saliencia (<i>keyness</i>) (fase 3)	
WordList (2ª parte)	Generación de listas de Información Mutua (IM) (fase 4)	Constituyen una primera aproximación a la naturaleza combinatoria de los términos: se identificarán gracias a los índices de IM que asignan un valor numérico al grado de ligazón de dos palabras que suelen co-ocurrir (colocarse) en el dominio analizado.
Concord (3ª parte)	Análisis de agrupaciones (<i>clusters</i>): listados poliléxicos (fase 5)	Corroboración o refutación mediante análisis en textos reales del carácter terminológico de las unidades previamente extraídas como términos potenciales. Identificación de agrupaciones léxicas (<i>clusters</i>), colocados, colocaciones y líneas de concordancia. Caracterización de los términos para su subsiguiente inclusión como entradas o subentradas de diccionario
	Análisis de colocados y colocaciones (fase 6)	
	Análisis de concordancias (fase 7)	

Tabla 5. Resumen del proceso de extracción y análisis terminológico llevado a cabo por WordSmith Tools (adaptado de Edo Marzá (2011)).

Tal y como se aprecia en la Tabla 5, con Wordlist el lexicógrafo realiza una primera aproximación al proceso de extracción semi-automática de términos al generar listados monoléxicos de frecuencia con las unidades del corpus. Un primer paso aconsejable entonces es la preselección de datos irrelevantes desde el punto de vista terminológico (fase 1) como pueden ser las palabras gramaticales y las unidades claramente “no terminológicas”. WST permite “filtrar” estas palabras gramaticales mediante la aplicación de listados que contienen esas palabras a eliminar/filtrar. La observación de los listados de frecuencia, generados esta vez sin palabras gramaticales (fase 2) es

normalmente el punto de partida para comenzar la selección y el análisis de los términos potenciales puesto que los listados libres de este tipo de “distractores” ahorran tiempo, energía y añaden fiabilidad a los resultados. El motivo básico para establecer este como el primer paso para la extracción terminológica no es otro que el hecho, generalmente aceptado, de que si observamos las primeras palabras generadas por un software de concordancias cualquiera en un listado monoléxico de frecuencia, veremos que, presumiblemente, los términos centrales de una especialidad pueden ser identificados. En palabras de Ahmad y Rogers (2001), estas unidades podrían calificarse como “términos madre” de una especialidad concreta y, como tales, engendran otros términos a través de los procesos de formación y combinación que sean válidos en la lengua objeto de estudio.

Sin embargo, también es cierto que la frecuencia es un criterio relativo, pues depende en gran medida de la naturaleza y equilibrio de los textos incluidos en el corpus, esto es, de la cantidad de textos sobre las distintas subáreas del dominio que se incluyan. En este sentido, como ya se había apuntado anteriormente, puede darse el caso de que una palabra altamente significativa aparezca sólo una vez o dos en el corpus porque pertenezca a un subdominio muy específico del que el corpus sólo contenga unas pocas muestras textuales. Por lo tanto, algunos términos que son interesantes para y representativos del campo de estudio es posible que aparezcan con una frecuencia muy baja o que directamente no aparezcan, fenómeno éste que suele originarse por una serie de circunstancias como la ya aludida naturaleza irregular del corpus, su falta de equilibrio o su tamaño inadecuado, de ahí la importancia de compilar un corpus fiable y cuidado. Es por ello que, para evitar descuidos o problemas derivados de fenómenos como el *hapax legomenon* (la palabra que aparece sólo una vez en un corpus determinado), las formas que aparecen solo una vez o con frecuencias a priori desdeñables deben ser también analizadas sin descartar de manera precipitada que puedan ser significativas desde un punto de vista terminológico.

Del mismo modo, también deben tratarse con cuidado las unidades léxicas que anteriormente hemos denominado “unidades claramente no terminológicas”. De acuerdo con la hipótesis de la unidad léxica será la situación comunicativa en la que se emplee dicha unidad la que activará en ella un valor especializado o no. Es por ello que una unidad

léxica que a priori consideraríamos no especializada como “naranja” puede ser un término en sí misma o ser elemento constituyente, como en el ejemplo (1), de dos términos polilexemáticos, “cáscara de naranja” o “piel de naranja”, que en el campo de la cerámica y en el contexto de nuestro corpus de estudio sí activan conocimiento especializado y han sido por tanto considerados unidades terminológicas:

1. *El revestimiento tipo **cáscara de naranja** se impresa con un rodillo especial, quedando una textura rugosa de **piel de naranja** y resulta ideal para cocinas, baños, quirófanos y todos aquellos espacios que requieren limpieza total y bajo mantenimiento.*

En consecuencia, tal y como ya habíamos apuntado, pese a la importancia de los cálculos de frecuencia, es cierto que esta no siempre es indicativa de un rol central en el dominio de especialidad y que puede verse alterada por un corpus de estudio inadecuado. Por ello, en nuestra propuesta abogamos por una extracción terminológica basada en sus primeros estadios tanto en criterios de frecuencia como de saliencia. La herramienta KeyWords (fase 3) ofrece un tipo de análisis complementario al ofrecido por WordList generando listados de palabras destacables o “salientes”, no “simplemente” ordenadas por frecuencia. El análisis conjunto de los resultados obtenidos con ambas herramientas ofrecerá un listado de términos potenciales basado en un doble criterio perfectamente complementario: la frecuencia (obtenida con WordList) y el carácter “saliente” de aquellas unidades léxicas que el programa ha identificado como infrecuentes al compararlas con un listado de frecuencia de un corpus de referencia general (Keywords). Esta combinación de datos de frecuencia con datos de saliencia dará lugar a un listado más refinado y fiable de unidades léxicas cuyo carácter potencialmente terminológico se encargará de corroborar o desmentir el resto de fases propuestas en el método.

En esta fase del trabajo se aconseja una primera aproximación a la naturaleza combinatoria y colocacional de los términos, clave en cualquier enfoque de base comunicativa. Por este motivo, en esta etapa del proceso, la combinación de los resultados obtenidos en el listado de términos monolexemáticos potenciales (basados en frecuencia y saliencia) se verá complementada por el análisis de listados de frecuencia poliléxicos o polilexemáticos que darán una primera idea al lexicógrafo especializado de las propiedades y tendencias

combinatorias o colocacionales de los términos. Especialmente interesantes resultan en esta fase del trabajo los listados poliléxicos de Información Mutua (IM) (fase 4), que “miden” la fuerza o el grado de dependencia mutua o ligazón de las unidades léxicas que forman términos compuestos y/o colocaciones. Estas unidades léxicas “mutuamente dependientes” tienden a “colocarse” (a aparecer formando colocaciones o sintagmas, esto es, en compañía habitual). Los listados poliléxicos, que también generaremos con WordList, están basados en la idea de que la palabra que aparece más frecuentemente con otra palabra que en cualquier otra posición dentro del corpus es vista como una combinación significativa y por lo tanto merece una consideración especial. Un índice de IM elevado es indicativo de que la asociación entre dos palabras es mayor de la que cabría esperar por puro azar o casualidad. Existe, sin embargo, un problema que el lexicógrafo especializado debe considerar desde un principio: el índice de IM no considera la frecuencia, de modo que puede inducir a error si una de las palabras presenta una frecuencia muy baja. Si una palabra de baja frecuencia ocurre sólo una vez con otra palabra, recibirá un valor de IM elevado a pesar de su baja frecuencia. Una vez más, este tipo de inconveniente puede compensarse complementándolo con otro tipo de análisis crucial en cualquier enfoque terminológico comunicativo, el de tipo textual, contextual y co-textual, es decir, el que permite al lexicógrafo analizar las unidades *in vivo* en el corpus.

Tal y como se resumía en la Tabla 5, en la última gran fase del proceso de extracción y análisis terminológico, Concord permitirá al lexicógrafo especializado la corroboración o refutación (mediante el análisis del uso real en el corpus) del carácter terminológico o no de las unidades previamente extraídas como términos potenciales. Para ello, se procederá a la identificación y el análisis de agrupaciones léxicas (*clusters*), colocados, colocaciones, y líneas de concordancia, lo que permitirá corroborar el carácter terminológico de los términos según su uso y facilitará su caracterización a nivel lingüístico para el subsiguiente relleno de las fichas terminológicas diseñadas y la consiguiente conversión de las mismas en entradas o subentradas de diccionario.

Así pues, la tercera gran parte de la fase 5 del método de trabajo propuesto consistirá en el análisis de agrupaciones léxicas (*clusters*) en forma de listados poliléxicos (fase 5). En palabras de YouJin (2009),

estas agrupaciones léxicas se definen como las secuencias de palabras más recurrentes en un registro y como tales deben ser consideradas como constructos lingüísticos básicos con funciones igualmente importantes para la construcción del discurso en diferentes lenguas. Con la aplicación Clusters, pueden generarse listados poliléxicos de 3, 4 ó 5 palabras (dependiendo de las especificaciones del lexicógrafo) y las unidades potencialmente terminológicas pueden ser por tanto analizadas desde una perspectiva cotextual y de tendencia a la agrupación. Por ejemplo, el análisis del término potencial *horno* en nuestro corpus con la herramienta Clusters y con unos parámetros predefinidos como los indicados en la Tabla 6, dan lugar a agrupaciones o *clusters* del tipo mostrado en dicha Tabla 6.

Parámetros para el análisis con WST	Agrupaciones léxicas resultantes o <i>clusters</i>
Palabras en el cluster: 4 a 4 Frecuencia mínima: 5 Horizontes: 4 a la derecha; 4 a la izquierda	...horno con atmósfera oxidante...//...horno con atmósfera reductora...//...atmósfera reducida del horno...//...horno con atmósfera reducida...//...

Tabla 6. Resultados obtenidos con Clusters según los parámetros establecidos.

Recomendamos aquí el análisis pormenorizado de los clusters en conjunción con los colocados y colocaciones de la fase 6 (ver Tabla 7). Esto nos permitirá avanzar identificando de manera más exhaustiva todos los posibles “patrones multi-palabra” fijos que pueden constituir términos polilexemáticos puesto que, junto con los resultados obtenidos con “Clusters”, la aplicación “Collocates” (fase 6) proporcionará al lexicógrafo una lista con todos los posibles colocados de la base (entendidos aquí en su acepción más simple y general, como palabras que aparecen en el contexto inmediato de dicha base) agrupados y mostrando la frecuencia que tienen en cada posición. La Tabla 7 muestra, de forma simplificada, una fila similar a la mostrada en la tabla de colocados generada por Concord para la base *horno* en la que se muestra la relación de esta base con su colocado frecuente *atmósfera*. Según los resultados obtenidos, corroboramos que efectivamente *horno* suele tener a *atmósfera* como colocado, especialmente en las posiciones segunda por la izquierda (I2) y segunda por la derecha, dando lugar, tal y como se muestra, a colocaciones (extraídas del corpus) del tipo mostrado en los ejemplos

(2) y (3). Estos resultados no hacen sino reforzar, en este caso, el tipo de resultados obtenidos con Clusters con respecto a la forma predominante en que se combinan *horno* y *atmósfera*. Obviamente, el análisis conjunto y completo de todos los resultados obtenidos en ambas aplicaciones (Clusters y Collocates) no puede ser mostrado aquí en su totalidad. Además, es indudable que el análisis comparativo de todos los resultados es una tarea costosa y difícil. No obstante, creemos que los ejemplos proporcionados abajo pueden ilustrar a pequeña escala lo que se pretende describir y conseguir.

Colocado	Total	Total izq. (I)	Total dcha. (D)	I3	I2	I1	Base <i>horno</i>	D1	D2	D3
ATMÓSFERA	45	30	15	2	18	0	0	1	6	3

Tabla 7. *Atmósfera* como colocado de *horno* según la aplicación Collocates.

2. "...en un *horno de atmósfera templada*..."
centro D1 D2

3. "...dependiendo de la *atmósfera del horno*..."
I2 I1 centro

Por último, Concord, y más concretamente su herramienta Concordance (fase 7) permite al terminógrafo observar las líneas de concordancia en las que los términos potenciales aparecen y analizar desde una perspectiva contextual más amplia la manera en que estas UCEs potenciales se usan en el discurso real propio de la especialidad. En esta séptima fase, la final según el método que aquí se propone, puede observarse como fundamento de este tipo de práctica lexicográfica el haber ido de lo concreto (la UCE en sí) a lo más amplio (la UCE en su contexto). Observar de este modo el funcionamiento de las unidades potencialmente terminológicas en el discurso es un paso fundamental para la correcta caracterización y presentación de los términos en los diccionarios especializados activos puesto que, como Gilquin (2002) apunta, su identificación es absolutamente fundamental para poder mostrar el aspecto comunicativo y real de la terminología de un dominio.

2.6 Procesamiento y almacenamiento de los datos

Una vez los términos han sido identificados tras el análisis pormenorizado de su uso en el corpus, el siguiente paso será almacenar

todos los datos en una base de datos (BD) terminológica, ya que el soporte automatizado permite, no sólo mantener la información permanentemente al día (con un coste razonable), sino también recuperar multidimensionalmente esa información de acuerdo con los intereses de cada trabajo (Cabré, 1993).

De acuerdo a los parámetros establecidos por Gómez y Vargas (2004), se aconseja que las BDs estén estructuradas de acuerdo a criterios consensuados y lógicos, de acuerdo a los usuarios y de acuerdo a la finalidad última de la compilación terminológica, que debe ser flexible y reflejar de manera precisa las relaciones entre las jerarquías de información, haciendo factibles la rápida introducción y recuperación de datos.

Es por ello que resulta crucial elegir un buen software o sistema de desarrollo y gestión de fichas terminológicas para poder elaborar y trabajar adecuadamente con esas fichas que vamos a elaborar. En el caso del diccionario de la industria cerámica, se eligió TermStar, un sistema de gestión de bases de datos terminológicas y multilingües orientadas al concepto. Desde estas BDs, el lexicógrafo especializado deberá, en primer lugar, diseñar una ficha terminológica acorde con la definición del trabajo hecha en la primera etapa y por tanto acorde con el tipo de producto que pretende desarrollar para, a continuación, proceder a rellenar los campos de dicha ficha terminológica con los datos extraídos del corpus y, si fuera necesario, con información adicional (normalmente de tipo técnico) procedente de las fuentes de consulta, con el fin de crear un repertorio terminológico por cada lengua de trabajo.

Es altamente aconsejable antes de proceder a rellenar las fichas terminológicas que el equipo de trabajo cree un protocolo escrito de elaboración de los registros o fichas terminológicas en el que se establezca de forma clara, sistemática y acorde a la finalidad de la obra tanto los estilos de definición y presentación de ejemplos, las distintas categorías a considerar en cada campo, la nomenclatura o codificación de las fuentes, etc. Por este motivo, se deberá elegir desde un principio un modelo claro de artículo o ficha terminológica, que en nuestro caso, y de acuerdo con la noción de diccionario activo, incluía (en TermStar) los campos denominados “categoría gramatical”, “campo conceptual”, “abreviación”, “contexto”, “fuente de contexto”, “fuente de datos”, “remisión”, “definición”, “fuente de definición”,

“sinónimo”, “principal” y “categoría”. Los campos señalados en la base de datos con un pequeño cuadrado verde al lado indicaban la existencia de un desplegable de categorías destinado a agilizar el trabajo y sistematizarlo. En el caso por ejemplo de las categorías “principal” y “categoría”, las codificaciones del desplegable constituían la clave para la posterior ordenación de los datos en entradas y/o subentradas dependientes de un término principal.

2.7 Revisión y resolución de problemas

Por último, las inconsistencias y los errores relativos a, por ejemplo, los campos (obligatorios) vacíos, categorías gramaticales erróneamente asignadas y errores ortográficos o de formato entre otros deberán ser corregidos. Dependiendo del sistema elegido en la fase anterior, esta revisión y resolución de los problemas detectados puede llevarse a cabo directamente en la BD o en otro programa. En nuestro caso, los datos de TermStar se exportaron a una tabla de Access en la que se revisaron y corrigieron, registro por registro, todos los posibles errores. En el caso de detectar problemas de difícil resolución por parte del terminólogo o su equipo directo de trabajo, se recomienda recurrir a los recursos documentales complementarios recogidos durante la compilación del corpus precisamente para ese fin o bien consultar de nuevo a alguno de los expertos designados al comienzo del proyecto.

Finalmente, en el caso del diccionario de la industria cerámica, se generó desde la tabla revisada de Access un documento *.rtf* (Rich Text Format) con la información final de las entradas. Este tipo de documento fue generado por ser compatible (Access no lo era) con el generador de diccionarios GenDic, utilizado en la última etapa del método.

2.8 Edición

Para finalizar, el diccionario deberá ser editado para lo que será necesario maquetar el trabajo, redactar el prólogo y las instrucciones de uso e incluir cualquier tabla, figura, índice o explicación que ayude al usuario a sacar el máximo partido a su diccionario.

En nuestro caso, el generador de diccionarios GenDic (desarrollado por la Dra. Chelo Vargas, experta en terminología y terminótica), fue el encargado de generar las entradas finales del diccionario con el formato mostrado en la Figura 8 (apartado 3) partiendo del documento en

formato *rtf* generado en la etapa anterior. GenDic generó las entradas ya maquetadas según lo que podríamos llamar un “formato entrada”, con lo en nuestro caso, una vez el diccionario esté totalmente concluido, sólo habrá que revisar de nuevo el conjunto final de entradas y añadir al principio y al final de la obra el prólogo y el resto de tablas o documentos aclarativos mencionados más arriba.

3. RESULTADOS

Creemos que el mejor resultado que podemos presentar en esta propuesta de fundamentos teórico-prácticos de la lexicografía especializada para la elaboración de diccionarios especializados es una entrada final del diccionario con el que venimos ilustrando todo el proceso. Es por ello que la Figura 8 muestra una entrada final del diccionario bilingüe de la industria cerámica y azulejera:

abrade V: ELABPROC desgastar, bruñir, corroer (metal), erosionar (roca), esmerilar, frotar, lijar, pulir, raspar ◊ *The sandpaper abraded the wood until it was smooth*; V. *abrase*. [Exp: **abradability** (PROPQUIM-FIS abrasividad; capacidad para desgastar por fricción elementos en contacto con un cuerpo y que guarda relación con la cantidad y dureza de la materia mineral que contiene dicho cuerpo ◊ *The low abradability of a pavement tile (0.06 g\sm) and staying power to sliding in winter are important features*), **abradability index** (PROPQUIM-FIS índice de abrasividad ◊ AS 1774.23.1-2001 *Refractories and refractory materials make the determination of their abradability index necessary*; V. *abrasion resistance*), **abradant** (MTRLAS abrasivo; material usado para abrasar -*abrade*-, molturar/triturar -*grind*-, desgastar -*wear down*- suavizar -*smooth*- o pulir -*polish*-tal que el esmeril -*emery*-, la arena -*sand*- o el vidrio en polvo -*powdered glass*-), **abrader** (INSTR V. *abraser*), **abrader** (INSTR abrasímetro; máquina de/para ensayos de abrasión), **abrading tool** (MAQTRANS/MAQELAB V. *abrasive tool*)].

Figura 8. Entrada final de *abrade* (en el diccionario bilingüe de la industria cerámica).

La entrada se presenta en su formato final, tras su generación automática en GenDic y muestra la entrada principal *abrade*¹ y las subentradas *abradability*, *abradability index*, *abradant*, *abrader* y *abrading tool*. En esta entrada observamos cómo el carácter comunicativo, social y activo del diccionario queda patente con la inclusión de contextos reales que ilustran el uso de los términos, del campo semántico al que pertenecen, de los diversos equivalentes en las lenguas de trabajo, de la categoría gramatical, y en general con la utilización de una microestructura densa en contenido pero creemos que útil para el correcto entendimiento y uso de los términos.

¹ Se ha optado por presentar una entrada con los términos en inglés para que así la mayor parte de la información lexicográfica de la entrada pueda leerse en español.

4. CONCLUSIONES

A lo largo del presente artículo hemos pretendido presentar de forma progresiva pero integrada y coherente una aproximación teórico-práctica al desarrollo de diccionarios de especialidad desde la lexicografía especializada. Con este fin, hemos comenzado analizando los fundamentos teóricos que debieran, a nuestro juicio, sustentar este enfoque, para lo cual se ha comenzado abordando el cambio producido del paradigma formal al funcional en lingüística, base de la actual concepción de la lengua como instrumento de uso y comunicación y clave a la hora de determinar la necesidad de estudiar la terminología en uso y, por tanto, en contexto. Siguiendo esta misma línea, se ha enfatizado la necesidad de estudiar los términos desde una perspectiva lingüístico-comunicativa, que aboga por aglutinar bajo un único marco teórico y bajo una misma etiqueta (lexicografía especializada) diferencias que no son reales entre la terminografía y la lexicografía. Las dicotomías terminografía-lexicografía y terminología-lexicología carecen de sentido según todo lo expuesto en este artículo por tratarse de marcos teóricos supuestamente distintos que, sin embargo, deben abordarse desde la lingüística por tratar en todos los casos la manera en que el conocimiento se vehicula mediante el uso de unidades léxicas. Los “términos” no son sino unidades de forma y contenido que activan su valor especializado dependiendo de la situación comunicativa en la que se dan. Es por ello que se insta aquí al lexicógrafo a adoptar la concepción y estudio de la terminología propuestos por la TCT, una teoría lingüística de base cognitiva y con un claro propósito comunicativo. La adopción de una perspectiva lingüística que asume que la terminología es lenguaje natural y no artificial –de modo que podemos afirmar que comparte las características del lenguaje general y no puede ser tratado como un fenómeno distinto (tal y como pretendían las teorías terminológicas más clásicas)– conduce al lexicógrafo especializado, necesariamente, a la adopción de un enfoque acorde con el hecho de que la terminología es esencialmente comunicativa, por lo que debe necesariamente ser real y fruto del uso y estudiarse y analizarse como tal. En consecuencia, la compilación de corpus textuales para el estudio de *in vivo* de la terminología se postula como el único medio posible para la elaboración de obras lexicográficas de calidad sustentadas en y orientadas al uso real.

Con toda la base teórica expuesta y que fundamenta la práctica lexicográfica especializada, se ha presentado también, si bien de manera sucinta, una sistematización metodológica para el desarrollo de diccionarios especializados activos, entendiendo por activo el hecho de que incluya datos suficientes, basados en el uso real, para la correcta utilización y comprensión de los términos, es decir, en el hecho de que promuevan y fomenten la comunicación de manera efectiva y acorde con las necesidades del usuario. En los últimos años se ha puesto de manifiesto la necesidad de prestar una mayor atención a los rasgos morfo-sintácticos y al comportamiento colocacional de los términos. Gracias al uso de una metodología ampliamente basada en el uso de las nuevas tecnologías (concretamente de las herramientas terminóticas) y de una buena fundamentación teórico-práctica, el lexicógrafo especializado puede, pues, desarrollar productos de calidad que satisfagan las necesidades reales de usuarios reales en situaciones comunicativas reales.

REFERENCIAS BIBLIOGRÁFICAS

- Adelstein, Andreína y María Teresa Cabré. 2000. Representación lexicográfica y terminográfica de las unidades terminológicas. Comunicación presentada en el *VII Simposio Iberoamericano de Terminología*, Lisboa, 14-17 noviembre 2000.
- Adelstein, Andreína. 2001. *Unidad léxica y valor especializado: Estado de la cuestión y observaciones sobre su representación*. Tesis de Doctorado. Universidad Pompeu Fabra, Barcelona.
- Ahmad, Khurshid, Martin, Willy, Hölter, Martin y Margaret Rogers. 1995. Special-ist Terms in General Language Dictionaries. University of Surrey Technical Report CS-95-14.
- Ahmad, Khurshid, y Margaret Rogers. 2001. Corpus Linguistics and Terminology Extraction, en Wright, Sue Ellen y Gerard Budin (eds.) *Handbook of terminology management*, vol. 2. Amsterdam/Filadelfia, John Benjamins, 725-760.
- Auger, Pierre y Louis-Jean Rousseau. 1987. *Metodologia de la recerca terminològica*, Sant Adrià de Besòs, Departament de Cultura de la Generalitat de Catalunya.
- Bergenholtz, Henning y Sven Tarp (eds.). 1995. *Manual of Specialized Lexicography: the preparation of specialised dictionaries*. Amsterdam/Philadelphia, John Benjamins Publishing Company.
- Biber, Douglas. 1990. Methodological issues regarding corpus-based analyses of linguistic variation, *Literary and Linguistic Computing*, 5, 257-269.
- Biber, Douglas, Conrad, Susan y Randi Reppen. 1998. *Corpus linguistics. Investigating language structure and use*, Cambridge, University Press.
- Cabré, M^a Teresa. 1993. *La Terminología: teoría, metodología, aplicaciones*, Barcelona, Editorial Antártida.

- Cabré, M^a Teresa. 1999. *La Terminología: representación y comunicación, elementos para una teoría de base comunicativa y otros artículos*. Barcelona: Universidad Pompeu Fabra.
- Cabré, M^a Teresa. 2000. Elements for a theory of terminology. Towards an alternative paradigm, *Terminology. International journal of theoretical and applied issues in specialised communication*, 6, 1: 35-57.
- Cabré, M^a Teresa y Andreína Adelstein. 2001. ¿Es la terminología lingüística aplicada?, en Carmen Muñoz (coord.) *Trabajos en lingüística aplicada*, Barcelona, Univerbook, S.L.: 87-393.
- Cabré, M^a Teresa y Rosa Estopà. 2006. Definición y redacción del plan de trabajo. *Máster online de terminología*. Universitat Pompeu Fabra.
- Codina, Luis. 2000. Evaluación de los recursos digitales en línea: conceptos, indicadores y métodos, *Revista española de documentación científica*, vol. 23, 1: 9-44.
- Edo Marzá, Nuria. 2009. *The specialised lexicographical approach: a step further in specialised dictionary making*. Peter Lang: Bern.
- Edo Marzá, Nuria. (En prensa). Corpus exploitation for terminographical purposes: a proposed term extraction process for bilingual specialised dictionary elaboration. *RæL-Revista electrónica de lingüística aplicada*, Volume 10.
- Gilquin Gaetanelle. 2002. Automatic retrieval of syntactic structures. The quest for the Holy Grail, *International Journal of Corpus Linguistics*, 7(2): 183-214.
- Gómez de Enterría, Josefa. 1998. Aproximación a los textos científicos. Una metodología activa, *Comunicación presentada en el XVI Congreso Nacional de Lingüística Aplicada*, Logroño – AESLA.
- Gómez González-Jover, Adelina y Chelo Vargas Sierra. 2004. Aspectos metodológicos para la elaboración de diccionarios especializados bilingües destinados al traductor, *El español, lengua de traducción. II congreso internacional*, Bruselas: ESLEtRA, 365-398.
- Hoffmann, Lothar. 1998. *Llenguatges d'especialitat*. Barcelona, Universitat Pompeu Fabra.
- Lerat, Pierre. 1995. *Las Lenguas especializadas*, Barcelona, Ariel.
- Posteguillo, Santiago y Jordi Piqué-Angordans. 2005. Computer Terminology: Developing an Active Bilingual English-Spanish Dictionary, en Gottlieb, Henrik, Mogensen, Jens, Eric y Anne Zettersten (eds.), *Symposium on lexicography XI. Proceedings of the eleventh international symposium on lexicography*, May 2-4, 2002 at the University of Copenhagen. Tübingen, Max Niemeyer Verlag, 453-473.
- Pitch, Heribert y Jennifer Draskau. 1985. *Terminology: an introduction*. Guilford: The University of Surrey.
- Rafel i Fontanals, Joaquim y Joan Soler i Bou. 2003. El procesamiento de corpus. La lingüística empírica, en M^a Antonia Martí Antonín (coord.): *Tecnologías del lenguaje*, Editorial UOC.
- Rey, Alain. 1979. *La terminologie: noms et notions*. Paris, Presses Universitaires de France.
- Rey, Alain. 1995. *Essays on Terminology*, Amsterdam/Philadelphia, John Benjamins.
- Rondeau, Guy. 1983. *Introduction à la terminologie*. Chicoutimi (Quebec), Gaétan Morin.

- Sánchez Pérez, Aquilino. 1995. Organización del Corpus Cumbre, en Aquilino Sánchez Pérez (coord.), *Corpus lingüístico del español contemporáneo. Fundamentos, metodología y aplicaciones*. Madrid, SGEL, 25-37.
- Santamaría Pérez, Isabel. 2006. La terminología: definición, funciones y aplicaciones. Biblioteca de recursos electrónicos de humanidades E-excellence. Liceus. Disponible en: <http://books.google.es/books?id=ZShn6fmIAmYC&pg=PA1&l>
- Sager, Juan Carlos, Dungworth, David y Peter F. McDonald. 1980. *English Special Languages: Principles and Practice in Science and Technology*, Wiesbaden, Oscar Brandstetter.
- Sinclair, John. 1996. *Preliminary recommendations on Corpus Typology*, EAGLES Document EAG-TCWG-CTYP/P. Disponible en: <http://www.ilc.cnr.it/EAGLES96/corpusstyp/corpusstyp.html>
- Tarp, Sven. 2005. La teoría de las funciones lexicográficas y su aplicación a la lexicografía especializada (1). Seminario impartido en la Universitat Pompeu Fabra.
- Tognini-Bonelli, Elena. 2001. *Corpus linguistics at work*, Amsterdam/Philadelphia, John Benjamins.
- Wüster, Eugene. 1979. *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie*, Wien: Springer (Versión española dirigida por M. T. Cabré Castellví: *Introducción a la teoría de la terminología y a la lexicografía terminológica*. Barcelona, Universitat Pompeu Fabra, 1998).
- YouJin, Kim. 2009. Korean lexical bundles in conversation and academic texts. *Corpora*, 4: 135-165.