

Recopilación y estructuración del vocabulario de especialidad en el *Nuevo Diccionario Histórico del Español* (RAE)

José Ramón Carriazo Ruiz

Marta Gómez Martínez

Centro Internacional de Investigación de la Lengua Española

The New Historical Dictionary of the Spanish Language is the latest work the Spanish Royal Academy has decided to undertake. In this lexical compilation, scientific and technical vocabulary will be taken into account and included, regardless of what other previous dictionaries, such as general language dictionaries, have done so far. For this purpose, a team of lexicographers in Cilengua (La Rioja) is studying the method to select, extract and tag the vocabulary used in specialized areas of knowledge. This presentation explains the different steps that will be followed in order to include the specialized terms and their history in the dictionary, such as the establishment of a representative corpus, the selection of terms or the introduction of subject matter labels.

1. Introducción

El Instituto Historia de la Lengua de Cilengua (Centro Internacional de Investigación de la Lengua Española), situado en San Millán de la Cogolla (La Rioja) colabora con la Real Academia Española en la elaboración del *Nuevo Diccionario Histórico del Español* (NDHE). En estas etapas iniciales, la organización y planificación de las tareas que configurarán la labor del equipo de lexicógrafos es fundamental para cumplir las metas de modo satisfactorio y en el periodo de tiempo establecido. A los investigadores del CiLengua nos ha correspondido la recopilación y estructuración del vocabulario de especialidad en sentido amplio¹, lo que Vitruvio denominaba *inconsuetus sermo*: el léxico oscuro o marcado frente al léxico común o no marcado.² Resulta muy difícil de cuantificar el número total de voces que supondrá esta aportación al NDHE, aunque es posible ya ofrecer una clasificación léxico-semántica de las distintas categorías de palabras esperables. La recopilación de los vocabularios profesionales históricos incluye, desde el principio, los términos científicos y técnicos (vocabularios científicos y terminologías históricas)³, pero también las palabras propias de las labores agropecuarias y oficios artesanales, denominaciones de objetos de la vida cotidiana (con las variantes dialectales características de estos campos léxicos)⁴, el léxico efímero (palabras sometidas a procesos de obsolescencia —arcaísmos, voces desusadas—⁵, campos léxicos con tendencia al neologismo en distintos grados y maneras —neologismos de la moda, el deporte, las costumbres, la divulgación

¹ Hemos decidido utilizar la etiqueta de *especialidad* o *especializado*, pues desde nuestro punto de vista ni *científico* ni *técnico* son denominaciones capaces de abarcar la realidad que pretendemos definir: aquellas unidades léxicas que recogen conceptos relacionados no solo con las áreas del conocimiento científico, técnico y profesional, sino también con las actividades de la vida cotidiana como la agricultura, la ganadería o la artesanía.

² Véase el concepto de *Neutral field of discourse*: “Core words do not normally allow us to identify from which field of discourse they have been taken. Thus the words galley, port [...] immediately recall nautical [...] contexts while corresponding items like *kitchen* [...] do not. This is of course not to say that for specific purposes such nautical lexis is not core in crucial sense. But for general purposes general items are such by virtue of their neutrality of field” (Carter 1998: 43).

³ Gómez Martínez (2006) y Carriazo Ruiz (2003).

⁴ La pervivencia de los arcaísmos técnicos en el uso dialectal es algo común a todas las variedades diatópicas de las costas de la Romania. (cf. Logue 2000: 335, Carriazo Ruiz 2007). Véase Gómez, Carriazo (en prensa).

⁵ Carriazo Ruiz (en prensa a) y Carriazo Ruiz, Giménez Eguíbar (en prensa).

científica—⁶ y el vocabulario familiar, el de mayor impronta cultural y sometido a variación no sólo diatópica, sino sobre todo diastrática (cursi, poético, literario, elevado, culto, familiar, vulgar, soez) y diafásica (infantil, juvenil, femenino/masculino).

Para el desarrollo de estas tareas se ha organizado una red de investigadores encargados del establecimiento de una propuesta de marcación lexicográfica que permita sacar el máximo provecho al diseño virtual del NDHE. A partir del núcleo central del vocabulario —*core vocabulary*— se han de definir distintos ejes de desvío o marcación posibles que dan lugar a colecciones de textos modulares de los que obtener los datos históricos pertinentes para cada área del vocabulario que debe recopilarse y estructurarse. De este modo, el NDHE partirá de un corpus central, cuyo fin es el reflejo lo más fiel posible del vocabulario histórico básico del español, al que se irán añadiendo corpus específicos para cada conjunto de vocabulario marcado (terminologías históricas, léxico tradicional, efímero, dialectal y familiar), que por fuerza ha de quedar excluido del corpus principal, fundamentalmente literario. En este trabajo analizaremos las tareas correspondientes al establecimiento de los corpus modulares pertinentes para la recopilación y estructuración de las terminologías históricas, los criterios para la distinción del léxico marcado, la investigación de la historia del vocabulario de naturaleza eminentemente icónica y el escollo que supone el léxico efímero.

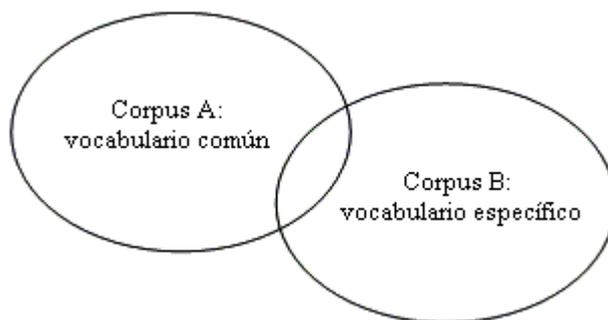
2. La estructura de corpus modulares abiertos

Partimos de un corpus modular, esto es, la base de trabajo del NDHE es un corpus complejo formado por varios tipos concretos de textos:

- Corpus A: textos de tipo general (no marcado)
- Corpus B: textos de especialidad (marcado)
- Corpus C: otros textos

Desde un punto de vista diacrónico es imprescindible seguir un método diferente para el estudio de las acepciones propias de un ámbito de especialidad frente a las que pertenecen a la lengua general. Para conocer el vocabulario de una lengua funcional histórica es necesario disponer de un corpus documental diseñado con el objetivo de disponer del léxico común, neutro o no marcado. Como cabe esperar, ese corpus dejará fuera un buen número de textos de especialidad de contenido demasiado especializado. Dado que la exhaustividad es necesaria y deseable en un diccionario histórico, es preciso crear un corpus específico para recopilar todos los documentos pertinentes para la recolección de los vocabularios propios de cada ámbito del saber y el análisis de su evolución histórica. Con el cotejo de estas dos colecciones de textos resultará más sencillo distinguir el léxico especializado del léxico común con criterios de tipo textual y contextual.

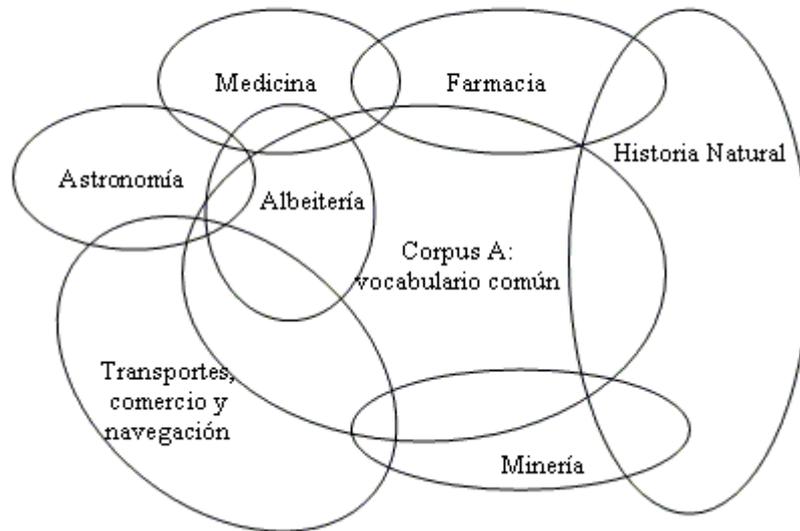
Esquemáticamente, el método para reconocer los términos específicos de cada área podría representarse mediante la intersección de dos grandes conjuntos de vocabulario:



Archivo "Corpus A y B"

⁶ Véase: "Una nueva herramienta lexicográfica: *CONCORDANTIA ORTEGIANA. Concordantia in José Ortega y Gasset opera omnia*", especialmente el apartado III: "El vocabulario de Ortega y Gasset a través de la *Concordantia ortegiana*". Algunos ejemplos: *amiba, de anastasio*. (Carriazo Ruiz en prensa b).

O más precisamente, cada una de las marcas correspondería a un corpúsculo que agrupe los textos con criterio temático (este esquema, como todo esquema, es reduccionista):



Archivo “Corpus modulares”

El vocabulario específico de distintos ámbitos profesionales presenta zonas de contacto con el vocabulario común y zonas de vocabulario compartido con otras áreas de actividad (medicina con albeitería y farmacia; historia natural con farmacia y minería; transportes, comercio y navegación con minería, astronomía y albeitería). Desde esta concepción, se pueden distinguir dos grados o tipos de vocabulario específico:

1) Léxico específico de una o varias actividades científicas pero no representado en el corpus de lengua común (vocabulario marcado en sentido estricto): sería el incluido en las partes de las áreas temáticas independientes, que no tienen contacto con la lengua común. Estos vocablos serán pocos, a pesar de la exhaustividad propuesta en la recolección, pero muy representativos, y pueden emplearse como guías para el establecimiento de las áreas temáticas de especialidad de vocabulario más característico (p. ej. *Leucoqueratina*).

2) Léxico compartido por la lengua común y los tecnolectos (vocabulario con una o varias acepciones técnicas): sería el situado en las zonas de intersección entre los círculos que agrupan el vocabulario específico (presente en los textos de una determinada especialidad) empleado también en la lengua común, en su acepción de especialidad o en otras (presente por tanto en los textos que reflejan el léxico común de una época). En muchos casos, se tratará de acepciones particulares con ecos científicos o técnicos (p. ej. *planta, animal, sol*).

3) Frente a estos dos grupos debe colocarse el léxico exclusivo de la lengua común —*core vocabulary*— (p. ej. *silla, husillo*).

Tanto los vocablos específicos exclusivos como los que comparten alguna acepción especializada (cuyo uso se documenta en el corpus de textos marcados temáticamente) y alguna vulgar (documentada en el corpus de textos no marcados temáticamente) han de recogerse de modo exhaustivo en un diccionario histórico. El NDHE debe dar cuenta, además, de los cambios de registro de las voces con el paso del tiempo, y del incremento de sus acepciones, como hechos básicos de la historia lingüística. Este sistema está diseñado pensando en ambos objetivos.

3. Criterios para la distinción del léxico marcado: criterio textual

El NDHE es un diccionario que dará cuenta del caudal léxico del español en su totalidad y reflejará sus cambios; por lo tanto, no cabe aquí aplicar los criterios utilizados en los diccionarios generales para la incorporación del léxico especializado pues, como recoge la bibliografía:

La primera dificultad surge al querer establecer criterios que permitan distinguir las voces generales de las especializadas. [...] El hecho de que haya términos especializados que penetran en el lenguaje común y al revés, palabras del lenguaje común que reciben un nuevo significado especializado, es prueba evidente de que la transición es fluida. (Gómez-Pablos 2002: 109).

Al parecer, el criterio más generalizado para la inclusión de tecnicismos en los diccionarios generales es el de la difusión. Por lo común, se procura incorporar al diccionario general aquellos vocablos científicos y técnicos que dejan de ser patrimonio exclusivo de círculos de especialistas para pasar a formar parte de la competencia del hablante culto o semiculto. (Alpizar 1990: 135).

En este sentido, Bertha M. Gutiérrez Rodilla (1993: 465) señalaba en su estudio sobre los términos relacionados con la medicina recogidos en el *Diccionario de Autoridades* que el problema básico que se le había presentado “para extraer las voces que tienen alguna relación con la medicina, es que no disponemos de un método que permita establecer cuáles pertenecen a este ámbito” y admite: “No he tenido más remedio que dejar a mi competencia como médico y hablante de español la responsabilidad en la elección de los términos. [...] No conozco en el campo de la lexicografía otro proceder diferente”.

El corpus modular que hemos planteado en el apartado anterior nos permitirá distinguir las acepciones de especialidad sin necesidad de acudir a la intuición: si una voz aparece en el corpus especializado con una acepción perteneciente a un ámbito del saber y no está en el corpus general, entonces parece claro que debe considerarse y marcarse como constituyente del léxico de especialidad. Si por el contrario estamos ante una unidad léxica que aparezca en ambos corpus, será el contexto el que determine su inclusión en el sistema especializado.

4. Método de investigación y exposición de la historia del vocabulario de naturaleza eminentemente icónica: el enfoque onomasiológico

La organización y recopilación de la información relevante para la explicación histórica del origen y desarrollo de estos tres tipos de palabras (términos específicos de un ámbito profesional, términos del léxico común con un uso particular en los textos de especialidad y palabras de parecida naturaleza referencial) ha de realizarse de acuerdo con los principios de la lexicografía onomasiológica. La marcación en los diccionarios de uso es un resto del método onomasiológico, común en la recopilación de enciclopedias y vocabularios de especialidad desde San Isidoro de Sevilla. El máximo desarrollo del sistema corresponde al método de Schuchardt “Wörter und Sachen” y se aplicó fundamentalmente en los atlas lingüísticos. En nuestro caso, la definición de las marcas guiará la recopilación del vocabulario, y puede servir también para la recuperación onomasiológica de la información en el producto final (como hace el *Trésor de la langue française*, al menos). Puesto que el NDHE se plantea, desde sus inicios, como un diccionario electrónico, los sistemas de marcación se convierten en elementos primordiales dentro del sistema de recuperación de información. A partir de una etiqueta, el usuario puede obtener todas aquellas acepciones marcadas; se olvida así la clásica búsqueda semasiológica de las compilaciones léxicas en papel. Además, mediante una opción de búsqueda compleja o combinada, el usuario será capaz de establecer sus propios criterios de selección de voces y acepciones marcadas diatópica, diastrática, diafásica o diatécnicamente.

5. El léxico efímero y los modelos de definición

La elaboración de un diccionario histórico comienza con la recolección del vocabulario nuclear proporcionado por el análisis lexicológico de un corpus principalmente literario, pues es la única fuente para el conocimiento del estándar antiguo. La lengua vulgar, familiar, profesional, regional y científica sólo puede conocerse a través del estudio de corpus modulares específicos diseñados con el objetivo de recopilar las palabras oscuras, el *inconsuetus sermo*, que por necesidad ha de quedar excluido del corpus literario nuclear. La lexicografía ha recurrido tradicionalmente a la marcación de esas voces mediante diversos sistemas de etiquetado más o menos inspirados en la descripción de los usos lingüísticos. Por su parte, la semántica histórica ha descrito de manera pormenorizada las diversas

particularidades que afectan al léxico marcado en su devenir diacrónico: cada tecnolecto es producto de la historia de la técnica correspondiente, las jergas y argots responden a la historia social del momento en que se emplearon, el eufemismo como resultado de los tabús culturales, la moda y el prestigio como causas de los procesos de obsolescencia léxica y neología, etcétera. Estas implicaciones, presentes también en el vocabulario central de las lenguas funcionales históricas, se agudizan en el caso de los lenguajes específicos, lo que sitúa como problemas centrales en el estudio histórico de éstos la cuestión del léxico efímero y de los modelos de definición.

Así como el *core vocabulary* suele ser más o menos estable en la historia de las lenguas, los vocabularios marcados son mucho más sensibles a los cambios de paradigma o culturales (el estándar escrito u oral sufre menos procesos de sustitución léxica, ampliación o restricción semántica, procesos metafóricos, como también es más estable y homogéneo en su gramática, debido a que es el resultado de un proceso de koinización o entandarización). En los lenguajes de especialidad a la variación terminológica se le une la variación conceptual; por un lado, la variación puede afectar a las denominaciones —variación denominativa— y así encontramos formas distintas y equivalentes, pues se refieren a un mismo concepto, debido a que a lo largo del tiempo han ido surgiendo diversas expresiones para recoger los contenidos de las distintas áreas de conocimiento especializado. Pero no sólo la lengua ha podido sufrir cambios, sino que debido al avance de las ciencias y de las técnicas, una rama del saber va variando sus preceptos por lo que la variación puede afectar a los conceptos —variación conceptual—. El tabú sociocultural provoca un rápido desgaste de los eufemismos, los cambios de paradigma científico-tecnológico acarrear cambios muy profundos en las terminologías, el léxico de la vida familiar y de las labores artesanales muestra una clara tendencia a la variación denominativa territorial y, en conjunto, los cambios socioculturales e históricos más efímeros y restringidos tienen siempre correlatos lingüísticos importantes para un diccionario histórico. El estudio de todos estos cambios nos llevará a la recopilación del léxico más efímero y de menor trascendencia para el estándar español mediante el diseño de corpus modulares de textos más o menos desechables: documentos de archivo, prensa o deposiciones de testigos en actas judiciales, por poner solo un ejemplo.

Al propio tiempo, se han de elaborar distintos modelos de definición en los que se incluyan los sistemas o, mejor, diasistemas, con sus diversas peculiaridades históricas, de modo que se refleje en el diccionario no sólo la historia del vocabulario estándar, sino también la historia de los vocabularios restringidos social, cultural o geográficamente, de tanta importancia para la comprensión de los textos del pasado. Es evidente que, tanto para el léxico específico, exclusivo y común, como para el vocabulario que designa los objetos de la vida cotidiana, resulta imprescindible en un diccionario histórico recurrir a la información enciclopédica. En primer lugar, debe identificarse y explicarse en todo caso el referente, bien mediante un dibujo (figuras geométricas, instrumentos de cocina), un esquema (partes de un barco, partes de la casa, del castillo, términos cosmográficos y astronómicos), una tabla (signos del zodiaco, días de la semana, meses del año), una descripción de la realidad (tipos de embarcaciones, elementos del vestir), una etiqueta (nombre científico de plantas y animales, símbolo de los compuestos químicos) o métodos similares. Además, el diccionario histórico ha de dar cuenta de los cambios que puedan afectar al significado con implicaciones lingüísticas y de los cambios de la marcación (trasvases entre áreas de especialidad, de la lengua marcada a la común, o viceversa). Para ello, así como para la marcación de las voces y la selección del léxico de especialidad, debe recurrirse al estudio material de la época, la historia de la ciencia y la técnica, de los oficios, artes y ámbitos profesionales en general, así como a la pequeña historia, para temas de la vida cotidiana como la cocina, la explotación agrícola, la transformación de los productos agropecuarios y áreas similares.

Referencias bibliográficas

- Alpizar Castillo, R. (1990). "El término científico y el Diccionario Académico". *Nueva Revista de Filología Española* 38 (1). 133-139.
- Carriazo Ruiz, J. R. (2003). *Tratados náuticos del Renacimiento. Literatura y lengua*. Valladolid: Consejería de Educación y Cultura de la Junta de Castilla y León - Universidad de Salamanca.
- Carriazo Ruiz, J. R. (2007). "Ictionimia y terminología marinera en la 'Recopilación de voces...' de El lenguaje popular de la Cantabria montañesa medieval de Adriano García Lomas". En Campos Souto, M.; Coteló García, R.; Pérez Pascual, J. I. (coords.). *Historia del léxico español*. A Coruña: Servizo de Publicacións - Universidade da Coruña. 39-48.
- Carriazo Ruiz, J. R.; Giménez Eguibar, P. (en prensa). "Procesos de sustitución léxica en el tecnolecto naval del Siglo de Oro: neología frente a obsolescencia". *Foro Hispánico*.
- Carriazo Ruiz, J. R. (en prensa a). "Organización de acepciones en terminografía histórica naval". *Actas del II Congreso Internacional de Lexicografía Hispánica*. (Alicante, 19-23 de septiembre de 2006).
- Carriazo Ruiz, J. R. (en prensa b). "Una nueva herramienta lexicográfica: *CONCORDANTIA ORTEGIANA*. *Concordantia in José Ortega y Gasset opera omnia*". *Revista de Lexicografía*.
- Carter, R. (1998). *Vocabulary. Applied Linguistic Perspectives*. New York: Routledge.
- Gómez Martínez, M. (2006). *Sacrobosco en castellano*. Salamanca: Universidad de Salamanca.
- Gómez Martínez, M.; Carriazo Ruiz, J. R. (en prensa). "La vid y el vino en el *Vocabulario Riojano* de Cesáreo Goicoechea". *II Congreso Internacional sobre el la Lengua de la Vid y el Vino y su Traducción* (Soria, 2-5 de abril de 2008).
- Gómez-Pablos, B. (2002). "El léxico especializado en el *Diccionario de Autoridades*". En Campos Souto, M.; Pérez Pascual, J. I. (eds.). *De Historia de la Lexicografía*. La Coruña: Toxosoutos. 107-118.
- Gutiérrez Rodilla, B. M. (1993). "Los términos relacionados con la medicina en el *Diccionario de Autoridades*". *BRAE* 73. 463-512.
- Logue, M. P. (2000). "Obsolete Occitan Loanwords of the French Language". En *Dissertation Abstracts International. Section A: The Humanities and Social Sciences*. University of Toronto.