

La importancia de las restricciones de selección en los diccionarios generales de español: ¿todavía “imprecisas”?

LAURA BALLESTEROS VALLADARES-ROCÍO JIMÉNEZ BRIONES
UNIVERSIDAD AUTÓNOMA DE MADRID

1. INTRODUCCIÓN

Las restricciones de selección han sido siempre un tema muy discutido desde que Katz y Fodor (1963) acuñaran este término dentro de la gramática generativa para referirse al mecanismo restrictivo cuya violación resultaba en la anomalía de la oración en cuestión. Desde entonces diferentes lingüistas las han tratado desde diferentes puntos de vista, pasando de posiciones más sintácticas (Chomsky 1965) a otras más semánticas, (Weinreich 1966) o (Coseriu 1967), hasta las últimas teorías cognitivas en las que la aceptabilidad de ciertas combinaciones léxicas sólo se puede explicar teniendo en cuenta el sistema de creencias de los hablantes y oyentes, los *frames*, (Taylor 1989: 92). Nos vamos a situar dentro de una de estas últimas corrientes funcionales-cognitivas (la Lexemática Funcional) y junto con los estudios estadísticos y de ocurrencia de contextos en un corpus (Corpus Cumbre) abordaremos las restricciones de selección en la dimensión léxica de los adjetivos incluidos bajo el término “impreciso”¹.

2. LA LEXEMÁTICA FUNCIONAL (LF) Y LA GRAMÁTICA DE CORPUS (GC)

2.1. Las restricciones de selección en LF

LF es el resultado de la integración de la Lexemática de Coseriu (Coseriu 1981) y del modelo de la Gramática Funcional (GF) de Dik (Dik 1978, 89). El Lexicón es la parte central de este marco teórico, considerado como una gramá-

1. Agradecemos al profesor Ramón Sarmiento que nos haya proporcionado los ejemplos del Corpus Cumbre.

tica en sí donde la palabra es el átomo central. Este lexicón se estructura en dos ejes: el paradigmático y el sintagmático. En relación al primero, los lexemas se organizan onomasiológicamente en campos semánticos siguiendo la Lexemática de Coseriu y las aportaciones de la LF del profesor Martín Mingorance. Para llegar al campo semántico se parte de las definiciones de esos lexemas de las que se extraen los rasgos hiperónimos compartidos por los posibles lexemas que formarán el campo semántico y aquellos que los diferencian entre sí. De esta manera tenemos una organización jerárquica con uno o dos archilexemas que dan nombre al campo semántico y una serie de dimensiones definidas por los rasgos diferenciadores entre estas unidades léxicas. Siguiendo este método de descomposición semántica hemos organizado nuestros adjetivos de la siguiente manera:

- (1) indefinido: 1. sin definir sus características o naturaleza 2. sin límites temporales
- impreciso: que carece de cualquier forma de precisión, claridad o definición.
- ambiguo: 1. impreciso al tener más de una interpretación.
2. impreciso al carecer de claridad o producir duda
- borroso: impreciso porque no se ve con claridad
- difuso: impreciso por su extensión y/o amplitud
- incierto: impreciso porque es dudoso o desconocido
- inexacto: incierto porque carece de precisión
- inconcreto: impreciso por su falta de concreción
- oscuro: impreciso al no comprenderse con facilidad (sentido negativo)

En cuanto al eje sintagmático, los lexemas aparecen con los patrones de complementación que ellos mismos subcategorizan. Toda esta información aparece bajo la forma del predicate frame (marco predicativo) (cf. Dik, 1989: 68-72), fórmula integradora bajo la cual aparece todo predicado en GF y que nosotras adaptamos para nuestras entradas tal y como aparece en (2):

- (2) *impreciso*A (x_1 , <prototíp. entidades que implican puntualidad o exactitud>) donde se especifica la siguiente información:

- a) La forma del predicado: *impreciso*
- b) Su categoría sintáctica: A(djetivo).
- c) La valencia cuantitativa: el número de argumentos que el predicado subcategoriza: (x_1)
- d) La valencia cualitativa: las funciones semánticas de cada argumento y sus restricciones de selección: (prototíp. entidades que implican pun-

tualidad o exactitud). Por lo tanto, para la LF las restricciones de selección son parte de las propiedades lingüísticas de cada predicado y como tal tienen que ser especificadas en su estructura semántica, es decir, en su marco predicativo correspondiente (Dik, 1989: 76). En cuanto a la forma en la que operan, se encuentran más cercanas a las transferencias de rasgos de Weinreich que a las restricciones de selección chomskianas, ya que su violación no es automáticamente equiparada con la agramaticalidad de esa predicación. Las restricciones de selección funcionalistas sólo identifican aquellos rasgos que son los *más compatibles* con la semántica de ese predicado en cuestión, de manera que si existe alguna desavenencia entre los rasgos especificados en el marco predicativo y aquellos del término a insertar se señala en la predicación y esto desencadenará una interpretación especial, la mayoría de las veces, metafórica. Respecto a la expresión de los rasgos semánticos, no seguimos la terminología restrictiva binaria (+/-) de las teorías de Coseriu o Geckeler. Es en este punto precisamente donde nos encontramos con el mayor obstáculo al tratar las restricciones de selección, ya que en casi todas las teorías semánticas o nada se dice sobre estos rasgos semánticos o se deja para que lo expliquen los filósofos o psicólogos (Lehrer, 1974:176). Así, el problema ha sido siempre el de elaborar una serie de rasgos lo suficientemente potentes como para captar nuestras intuiciones sobre aquellas combinaciones léxicas posibles o no en nuestro idioma. Es muy fácil encontrar una palabra y ponerla entre corchetes y llamarla rasgo semántico, pero hay que explicitar las razones para cada uno de ellos (Lehrer 1974: 178). Por supuesto, no está en nuestras manos tal empresa; lo que nosotras pretendemos es algo más restringido y limitado al ambiente de un diccionario, así somos conscientes de las limitaciones de los rasgos usados en nuestras definiciones.

2. 2. Gramática de Corpus

Siguiendo las palabras del profesor Sánchez (1995:27):

Un corpus lingüístico es un conjunto de datos lingüísticos, sistematizados según determinados criterios, suficientemente extensos en amplitud y profundidad de manera que sean representativos del total, o de un ámbito, del uso lingüístico 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.

A partir de la creación de corpus lingüísticos con estas características, se pueden realizar estudios comparativos o contrastivos de la lengua. Ha sido una constante histórica escribir gramáticas de la lengua para ayudar a comprender la lengua a nativos y extranjeros. Hasta ahora, estas gramáticas eran descritas a partir de la intuición que el gramático, como hablante de la lengua, poseía sobre ella. El problema es que dicha noción de la lengua podía no coincidir con la realidad

lingüística, es decir, con el uso que el resto de los hablantes hacen de la lengua. Frente a esta gramática, surgen los corpus, que sugieren que: en lugar de partir de las reglas para buscar ejemplos, partamos del mismo uso para crear las reglas que expliquen la lengua. Nosotras hemos utilizado el corpus de la lengua española *Cumbre* (SGEL) como fuente de información para nuestra investigación. En primer lugar buscamos en el corpus cada uno de los adjetivos que forman esta pequeña dimensión. Una vez que tuvimos todas las muestras de lengua del corpus, comenzamos a trillar los significados que no se ajustaban a nuestros intereses específicos. Después analizamos cada caso en profundidad, teniendo en cuenta su contexto oracional y textual. Centramos nuestra atención en estudiar los rasgos semánticos de las palabras que acompañan a este adjetivo, sin descuidar los complementos que el mismo adjetivo requiere ni los casos metafóricos. En cuanto a las ventajas que nos proporciona el uso del corpus, en primer lugar, es más objetivo que la propia intuición como fuente de datos, ya que se trata de un número importante de palabras de la lengua en su contexto de uso. En segundo lugar, dicha amplitud de ocurrencias nos permite descubrir o darnos cuenta de ejemplos de uso en los que no habíamos pensado. En tercer lugar, los útiles que el corpus informatizado posee (KWIC, concordancias, etc.) nos permiten tener una amplia visión de nuestra lengua desde diferentes puntos de vista e incluso realizar trabajos estadísticos de ocurrencias. Todo esto nos ha ayudado a realizar un trabajo objetivo. Nuestra investigación nos ha llevado a una conclusión general: cada ejemplo justifica y desequilibra a la vez el sistema lingüístico, por lo tanto, el sistema y el ejemplo son dos cuestiones complementarias. Lo que llamamos lengua (sistema) y lo que llamamos uso (ejemplo) son dos observaciones de un mismo fenómeno desde distinto punto de vista. También hay desventajas del uso de corpus como base del estudio científico. Primeramente, el corpus es un conjunto finito de elementos en el que se excluyen las construcciones no posibles del sistema, puesto que no aparecen documentadas. Por lo tanto, el corpus como fuente de investigación impide utilizar el dato negativo. En segundo lugar, al ser limitado, es imposible que recoja todas las ocurrencias posibles de un elemento de la lengua, por lo tanto siempre quedará como una herramienta representativa de ésta. A pesar de estos inconvenientes, hoy por hoy no se puede prescindir del uso de corpus en el estudio de la lengua ya que proporciona una descripción gramatical de las variedades de uso lingüístico actual, tanto del código oral como del escrito.

3. CODIFICACIÓN DE LAS RESTRICCIONES DE SELECCIÓN: MARCO PREDICATIVO Y APLICACIONES AL DICCIONARIO CD-ROM

A continuación les presentamos un ejemplo de lo que sería nuestra entrada en un posible diccionario en formato CD-Rom:

- (3) Borroso: impreciso porque no se ve con claridad
borroso_A (x_1 :<prototíp. entidades que se ven>)
Ej: Y su figura era borrosa, como si entre ella y él se interpusiera la lluvia borroso_A (Metaf. e_1 :<prototíp. estados de cosas>)
Ej: Los estímulos de la carne tenían una sugerencia muy borrosa e imprecisa

Usamos un marco predicativo para cada uso del adjetivo en cuestión, junto con su correspondiente ejemplo. Véase la entrada dada por los cuatro diccionarios de español de este mismo término: la información de las restricciones de selección es bastante pobre y en algunos casos inexistente.

4. CONCLUSIÓN

En este trabajo hemos presentado el complicado concepto de las restricciones de selección dentro de un nuevo entorno: el diccionárico. Se ha mostrado cómo la combinación de dos teorías a priori sin mucha conexión, como son la LF y la GC, son un instrumento potente capaz de mejorar el novísimo mundo de los diccionarios informatizados, y, cómo no, de codificar nuestras intuiciones sobre lo que sí se puede decir o no en nuestro idioma.

BIBLIOGRAFÍA

- DIK, S.C. (1989): *The Theory of Functional Grammar. Part I: The Structure of the Clause*. Dordrecht: Foris Publications.
- FABER, P y R. MAIRAL. (1994): "Methodological Underpinnings for the Construction of a Functional Lexicological Model" en *Miscelánea: A Journal of English and American Studies*, 15, pp. 193-217.
- LEHER, A. (1974): *Semantic Fields and Lexical Structure*. Amsterdam: North-Holland Publishing Company.
- LIPKA, L. (1990): *An Outline of English Lexicology: lexical structure, word semantics, and word-formation*. Tübingen: Niemeyer.

- MARTÍN MINGORANCE, L. (1990): "Functional Grammar and Lexematics" en J. Tomaszczyk y B. Lewandowska-Tomaszczyk (eds.), *Meaning and Lexicography*. Amsterdam/Philadelphia: John Benjamins.
- SÁNCHEZ, A. (1995): Cumbre. *Corpus lingüístico del español contemporáneo. Fundamentos, metodología y aplicaciones*. Madrid: SGEL
- SIEWIERSKA, A. (1991): *Functional Grammar*. London: Routledge.
- TAYLOR, J. R. (1989): *Linguistic Categorization. Prototypes in Linguistic Theory*. Oxford: Clarendon Press.
- ZERNIK, U. (1991): *Lexical Acquisition: Exploiting On-line Resources to build a lexicon*, LEA. London.

Diccionarios

Clave. Diccionario de Uso del Español Actual. 1996. Madrid: SM

D.R.A.E. 1992.

Diccionario de Uso del Español . 1966. Gredos.(M^a Moliner)

Diccionario Salamanca de la Lengua Española. 1997. Universidad de Salamanca: Santillana.