

INSTITUT D'ESTUDIS CATALANS, *Diccionari descriptiu de la llengua catalana*, Institut d'Estudis Catalans, Barcelona, [2016].

MARÍA PAZ BATTANER ARIAS
Universitat Pompeu Fabra
Real Academia Española
paz.battaner@upf.edu

La total puesta a punto en consulta pública del *Diccionari descriptiu de la llengua catalana* (DDL) del Institut d'Estudis Catalans a finales de 2016 representa una novedad digna de señalarse en la lexicografía de las lenguas occidentales y, muy en concreto, en las lenguas románicas¹. Además, la transparencia de información sobre su estructura y las decisiones adoptadas en él que ofrece el sitio son notables.

El diccionario ha tenido un largo período de elaboración, cuya primera fecha se remonta a 1985, y ha requerido previamente la constitución de un corpus del catalán contemporáneo, fuente única de su nomenclatura, y una base de datos para su proceso de redacción. A lo largo de estos treinta años ha pasado por diversas consultas sobre lo que se iba redactando, aunque no se hubiera terminado de elaborar, lo que confirma el grado de transparencia informativa.

Como diccionario institucional que es, ofrece una "Presentación" en la que se explica en tono divulgativo su carácter lexicográfico, como todo diccionario general exige según el público al que va dirigido. A continuación, la "Guía para su utilización" explica las decisiones que se han tomado para sus prestaciones lexicográficas, para la información léxica que atiende y la complejidad con que las trata. Las características lexicográficas van en paralelo con la concepción del vocabulario catalán que lo informa, pues todo diccionario, explícita o implícitamente, responde a una teoría del léxico. Así, "Presentación" y "Guía", junto a la consulta de las "Entradas", la parte tercera y central del diccionario, constituyen la superestructura del DDL. La consulta se ofrece en dos versiones, la sencilla y la completa, que el usuario puede escoger en función de lo que quiera encontrar. La consulta más habitual será la denominada "sencilla", esto es, por lema.

Desde el formato electrónico, su macroestructura se diluye; el número de entradas y el tipo de lemas descritos son difíciles de calibrar desde la versión digital. Si en el formato en papel y en orden alfabético, un diccionario muestra fácilmente su grado de atención

¹ El diccionario puede consultarse en <http://dcc.iec.cat/ddl/index.asp>.

en cuanto al vocabulario descrito, no ocurre así en uno ofrecido en la red. Este tipo de formato oculta datos que están presentes en los diccionarios impresos y, de forma aún más tajante, cuando se presentan en cederrón con programas de consulta incorporados que permiten cuantificar mucha información. Por comunicación personal con su coordinador, Joan Soler, que agradezco mucho, puedo aportar que, a fecha de agosto de 2017, el DDLC estaba formado por 109 998 entradas, de las cuales 78 305 son entradas principales que desarrollan toda la estructura de los artículos y 31 693, entradas secundarias, que sólo constan de información remisiva a un artículo principal. Este número puede variar mínimamente por alguna incorporación o corrección efectuadas.

Las características y la calidad de la información que contiene, sin embargo, se observan fácilmente en la organización de su microestructura, que atiende a cuatro facetas diferenciadas del componente léxico de las lenguas románicas y que seguiré para reseñarlo.

La primera información es cuantitativa, en frecuencias absolutas, y cualitativa, de categorización gramatical de cada unidad léxica. Cada lema está atendido solamente bajo una categoría gramatical, entre otras, que pueda actualizarse. Ningún cambio de categoría aparece dentro de ningún artículo; las voces que pueden presentarse bajo dos categorías gramaticales obtienen tantos artículos como categorías, o bien se consideran derivados regulares y aparecen sucintamente explicadas al final del artículo de la base de derivación. Un ejemplo bien representativo de este tratamiento se observa en *vell, vella* (esp. *viejo, vieja*), que recibe dos artículos por cada una de sus categorías gramaticales en el discurso, como adjetivo y como sustantivo masculino y femenino.

Otra información previa sobre el lema que anuncia la entrada es la estimación de frecuencia de uso de la voz señalada con cuatro círculos, que llega a tener cuatro en negro en las muy usuales y a ir disminuyendo hasta presentar cuatro en blanco. Tras esta última información, y de forma esquematizada, pero cuantificada ahora con precisión, aparece también el monto de cada forma flexiva del lema en sustantivos y verbos; los adjetivos no ofrecen estos datos y tampoco los pronombres. Para los sustantivos, los datos cuantitativos se dan por género, si tienen esta moción, y por número; para los verbos, se recogen las formas personales (encontradas en los modos indicativo, subjuntivo e imperativo) y las no personales (infinitivo, gerundio y participio), además de las formas de los tiempos simples y las de las personas gramaticales. Veamos los datos cuantitativos del verbo *disposar* (esp. *disponer*):



F	65	35							
NP	66	12	22						
M	88	9	4						
T	57	19	9	3	1	6	3	4	
P	4	2	60	11	2	21			

Estos datos cuantificados del corpus pueden advertir al consultante de muchos rasgos de la voz explorada y requieren para su interpretación un alto conocimiento de las características de la lengua. En otro tipo de información con este mismo grado de detalle se constata que el DDLIC no es un diccionario para un público general que tenga sólo problemas de descodificación, sino para un analista o para quien quiera producir un texto y se enfrente, por ejemplo, ante el dilema de escoger entre una construcción en forma personal y otra en impersonal según el modo en subjuntivo o en indicativo. Un paso más sería ofrecer agrupadas las concordancias del corpus que están contabilizadas, el cual podría darse en un futuro.

La información sintáctico-semántica y ejemplos llenan la segunda parte del artículo. Las acepciones responden a las explicaciones semánticas y sintácticas de la forma simple del lema o de las formas complejas en que se incluyan. El DDLIC distingue entre las que se explican ya sea propiamente o ya por remisión a otro lema. Durante su larga gestación y muestra pública, las remisiones no se hacían patentes si no estaban redactadas, de manera que la consulta ofrecía desde sus principios datos ya fijos, aunque no estuvieran completos.

En cada acepción que lo requiere se inserta el patrón sintáctico esquematizado en que el lema se muestra con su valor correspondiente y en que se atiende a datos morfológicos, argumentales o puramente sintácticos. Esta información, que antecede a la definición, es otro elemento más que precisa de conocimientos lingüísticos del consultante. Las variedades de los patrones exigen interpretación ajustada, de aquí que también se especifiquen los elementos sintáctico-semánticos del patrón con una etiqueta de tipo ontológico². Los patrones se ofrecen en cada acepción de sustantivos, adjetivos, verbos y adverbios, y sus elementos, caracterizados por subíndices, se incorporan a la definición que les sigue, como se ve en el ejemplo sencillo del adjetivo *dispost, disposta* (esp. *dispuesto*) en su segunda acepción³:

² El ejemplo de categorización semántica que ofrece la “Guía” para *consciència* es “[humà, món, activitat intel·lectual, realitat, fet]”, dado que se interpreta como sustantivo no contable y sin referente físico.

³ Casos de descripción casi exhaustiva son, por ejemplo, los adjetivos, deverbales también, *endarrerit, retardat, atrassat*, que reciben explicaciones de uso, según los sustantivos con los que vayan y los verbos con que se atribuyan.

2. [N₁ (és) **ADJ**] (N₁[humà])

Se muestran así las dos construcciones del adjetivo, en atribución con el verbo *ser* (cat. *esser*) y en grupo nominal; en las dos construcciones, el sustantivo N₁ ha de cumplir el rasgo ‘humano’. El grafismo que se utiliza para el esquema de estos patrones es complejo: abreviaturas, paréntesis de dos tipos u otros signos para describir las diferentes construcciones de la voz. La definición de ese mismo adjetivo, “[Algú]₁ amb disposició⁶”, sigue los dos componentes necesarios para caracterizarla: los elementos semánticos y la estructura del patrón sintáctico que se ha mostrado anteriormente. La relación entre el patrón y la definición se establece explícitamente utilizando los subíndices que remiten a los elementos del patrón sintáctico, los cuales presentan el mismo subíndice (en este sencillo caso sólo el sustantivo lleva el subíndice que lo identifica como ‘humano’).

Por la atención precisa que reciben los elementos argumentales, muchas definiciones resultan sinónimas exclusivamente, aunque, como van acompañadas de los esquemas sintácticos y argumentales, logran ser muy precisas. Sólo en las acepciones que requieren construcciones con varios argumentos, la información semántica puede presentar alguna dificultad de lectura. La que se recoge de las identificaciones semánticas de los argumentos daría para realizar un interesante trabajo a partir del DDLc. Por ejemplo, una subacepción de *disposar*, que lleva aclaraciones sintácticas (con subíndices) y semánticas (entre corchetes para los tres argumentos), sería el esquema representativo de la voz en la cita siguiente: “Formulada la denúncia, el Tribunal *disposa* el convenient per a la pràctica de les diligències i actuacions encaminades a preparar el judici... [Rubió i Tudurí (1937): 34, p. 31]”:

2b. [N₁ **V** N₂ (per a N₃/per a Vinf₃)] (N₁[humà]; N₂[cosa, lloc, cosa necessària]; N₃[finalitat]) [Algú]₁ preparar [una cosa, un lloc, allò que és necessari]₂ [per a una finalitat]₃.

En general, la definición de las unidades léxicas descriptivas toma la redacción de una paráfrasis substitutoria con género (*genus*) categorizador y diferencia específica; en las definiciones substitutorias queda resaltado en color azul el género bajo el que se clasifica semánticamente la voz; en el ejemplo anterior, éste sería *preparar*, única parte de la definición, pues las otras son argumentos que coinciden con lo necesario para la actualización de *disposar* en el uso que se describe: un sustantivo de rasgo ‘humano’, un segundo sustantivo con las características de ‘cosa’, ‘lugar’, ‘algo necesario’, y un tercer sustantivo potestativo ‘finalidad’ con la que se hace.

No quedan proscritas las definiciones metalingüísticas, que también se ofrecen en alguna acepción de lemas que tienen usos marca-

dos pragmáticamente. Por ejemplo, los valores sociales del sustantivo *gràcia*, en su más frecuente uso en plural, como ocurre en español, primera de sus acepciones:

1a. [N_{INCOMPT}_{pl} (a N_i) (per N₂/per V_{inf}_{2i}/de N₂/de V_{inf}_{2i})] (N₁[humà]; N₂[benefici]) Denota gratitud a l'interlocutor per un benefici que se n'ha rebut. ⇒ *gràcies per tot, gràcies pel favor, gràcies pels beneficis, gràcies pel consell...*

Podría decirse que el diccionario se caracteriza por una descripción meticulosa de los patrones sintáctico-semánticos de los lemas, lo que determina las acepciones o subacepciones, y por una simplificación de las definiciones, que alguna vez queda eclipsada por los patrones que lleva y por los largos ejemplos de corpus que la ilustran.

Por todo lo anterior, la orientación codificadora es muy importante en el DDLIC, a tal punto que parecería predominar sobre la descodificadora, o probablemente esta percepción se deba a que la función codificadora suele atenderse menos en los diccionarios convencionales y, en cambio, destaca en el que reseñamos. De todas maneras, para ser utilizada, precisa de una formación lingüística avanzada.

La ordenación de las acepciones atiende a los principios de frecuencia y agrupación lógica o semántica, aunque a menudo parecen ofrecidas en orden inverso: agrupación lógica y frecuencia. El diccionario distingue entre verdaderas acepciones de valor semántico nuclear (numeradas) y, dentro de ellas, entre posibles subacepciones (diferenciadas con letras minúsculas). Por un lado, se analizan con detenimiento los matices semánticos de las palabras descritas en función de los patrones en que se muestran y, por otro, se ahorra en información para algunos derivados. Veamos ahora un esquema de la ordenación de acepciones del sustantivo, derivado del mismo verbo que venimos utilizando, pero con suficiente entidad propia para llevar artículo, *disposició* (esp. *disposición*):

1. Remisión a *disposar* [esp. *disponer*; marca así la nominalización del verbo]

Contable

2. (efecto, objeto, 'norma', resultado, y por eso contable)
 - 2a Precepto
 - 2b Aspecto de
 - 3c Parte de un documento
 - 2d Indicación

No contable

3. (efecto, sujeto paciente)
4. 3a Manera de estar dispuesto
 - 3b Manera de estar organizado
5. Poder de disponer [esp. *libre disposición*]

6. (estado)
 - 5a Estado de espíritu
 - 5b Estado de espíritu condicionado
 - 5c Estado físico
 - 5d Estado adecuado [esp. *en disposición de*]
 (No contable, cualidad, inclinación, tendencia a)
7. Voluntad para hacer
8. Aptitud
9. Predisposición²
10. Predisposición³

Dentro de cada acepción numerada⁴, e introducidas por una flecha, se ofrecen coapariciones características, del tipo de *paisaje pintoresco*, llamadas combinaciones. En el ejemplo anterior, se encuentra en la acepción 5b, ⇒ ~ *d'ànim* // ~ interior (esp. *disposición de ánimo* // *disposición interior*). En ocasiones, con otro tipo de signo, se muestran los cambios de categoría sintáctica que el lema puede tomar a los que se trata como derivados: el participio verbal en adjetivo o el infinitivo en sustantivo; por ejemplo, en el verbo *filar* (esp. *tejer*), el participio es tratado como adjetivo y se muestra con el signo ⇒ *adj.*, junto al ejemplo: “Són rares, però encara hi ha cases selvatanes que tenen envejables paraments de llit i de taula fets amb teles de lli *filades a casa* [Fort (1961): 39, p. 54]”.

No ocurre lo mismo con adjetivos que se muestran como sustantivos ni con infinitivos que se muestran como sustantivos, pues se consideran categorías diferentes y reciben, cada uno de ellos, artículo particular; caso de *vell* (cf. *supra*) o *retòrica*. En las palabras funcionales, la ordenación responde a sus usos y al matiz que se les atribuye; un ejemplo en el que se observa fácilmente esta solución es el tratamiento del posesivo *teu*, *teva* o *teua* (esp. *tuyo*, *tuya*).

La información sobre los derivados regulares, la información de variantes y los tratamientos de la voz en otros diccionarios del catalán también se atienden en el *Diccionari descriptiu*.

La “Guía” es muy útil para entender el tratamiento de las voces derivadas tal como se encontrarán en la consulta. Todos los derivados, sean regulares o no, están listados en la macroestructura, de manera que en la consulta electrónica aparecen fácilmente. Sin embargo, los derivados regulares, morfológica y semánticamente sólo quedan explicados al final del artículo de la forma base correspondiente, con la indicación de la acepción con la que coinciden por significado. Los derivados cuyo valor semántico y su construcción gramatical no sean regulares llevan artículo propio y se tratan exactamente como las unidades léxicas presentadas en una entrada con lema. Puede

⁴ Cabe señalar que las subacepciones 5b y 5d y las acepciones 8 y 9 del esquema van marcadas como no normativas en el catalán contemporáneo.

ser el caso del adjetivo *concís*, *concisa*, con sus acepciones propias y la inclusión de *concisament* y *concisió*, que aparecen descritos al final del adjetivo base, *concís*, con una indicación numérica de su construcción morfológica regular de derivación para adverbios y sustantivos derivados de adjetivos, pero sin glosa semántica.

Entre otros tipos de información, el diccionario ofrece variantes formales: *a*) de extracción dialectal o sociolectal, *b*) apreciativas (formas superlativas o comparativas) y *c*) flexivas que no coincidan con los paradigmas normativos. Así mismo, se encuentran muchas advertencias de formas de lemas y de usos semánticos o gramaticales no normativos, entre ellas, muchas formas de influencia del castellano en la lengua catalana, que son difíciles de detectar para los no filólogos. Vienen recogidas, remitidas y precedidas de un pequeño círculo negro que avisa de su consideración puramente normativa con respecto al *Diccionari del Institut d'Estudis Catalans (DIEC2)*, heredero del diccionario dirigido por Pompeu Fabra de 1932 y tenido por el punto de arranque que determina la norma del catalán contemporáneo. La confirmación que surge tras esta consideración es que un diccionario descriptivo no puede por menos de subordinarse a la normativa. El catalán ha recibido y recibe mucha atención normativa debido a su historia y a su larga ausencia en la escuela. Por lo demás, muy cuidadosas son las variantes de las locuciones, unidades léxicas que, aun cuando tengan la característica de fijación, no es extraño que con frecuencia presenten variaciones formales.

En atención a la historia de la lengua catalana, el diccionario descriptivo recoge, cuando es necesario, datos que en otros diccionarios anteriores del catalán aparecen como complementarios. Este recorrido por repertorios del catalán tenidos en gran estima convierte en parte al DDLC en un tesoro lexicográfico abreviado y útil para los sensibilizados con la historia de la lengua.

Interesa en esta reseña el carácter lexicográfico del diccionario, no la perspectiva interna de la descripción lingüística del catalán, aunque sean estos dos aspectos obligatoriamente confluyentes en una obra lexicográfica. Por ello, en lo que resta me dispongo a hacer ciertas apreciaciones generales sobre lexicografía digital.

El hecho de haberse planeado el DDLC con anterioridad a la *Gramática* (2016) normativa de la misma institución (GIEC) y a la nueva *Ortografía* (2017), y de haber tenido un proceso de elaboración más largo, puede explicar que se encuentren algunos desajustes entre estas dos obras y el presente diccionario descriptivo (1985-2016), además de que también explicaría suficientemente la diferencia entre los adjetivos *normativo* y *descriptivo*. Casos no normativos, comentados en reseñas hechas sobre la *Gramática del Institut d'Estudis Catalans*⁵,

⁵ Agradezco a mi compañero del Departament de Traducció i Ciències del Llen-

podrían servir de ejemplo; el muy usual apelativo *papa*, que el DDLC remite a *papá*, donde se lo define sencillamente como ‘padre’, sin ninguna otra información sobre su uso social ni sobre su origen y, por tanto, sin marcar como no normativo, entra en contradicción con la valoración de *papa* como normativo frente a *papá* como voz marcada. El adjetivo descriptivo es menos comprometido que el normativo en obras institucionales como éstas; lo normativo exige una posición de rigor o de tolerancia que siempre puede ser discutida o rechazada, en tanto que lo descriptivo requiere sólo la recogida y el análisis adecuado.

El aspecto formal en pantalla es crucial en un diccionario consultable en la red. Para ofrecer la muy rica información del diccionario, el grafismo se sirve de algún color: el género de la definición va en azul y la voz ejemplificada en rojo. También hay una leve sombra amarilla para dar entre corchetes interpretaciones semánticas del lema en esa acepción. Lo demás se confía a la diversidad en el tipo de letra: puntos, fuentes, signos sustitutorios del lema como ~, etc., todo en negro, con lo que el aspecto en pantalla se aproxima al de un diccionario en papel. Como por el momento no se ha previsto adelantar en un sucinto menú la estructura del artículo consultado, a veces resulta arduo ver en pantalla toda la riqueza de cada unidad léxica o locución anteriormente expuesta en el diccionario; se detecta todavía una inercia del formato en papel; por tanto, un esquema sería de gran ayuda en los artículos altamente polisémicos.

El diccionario se ofrece con dos tipos de consulta, como se ha dicho; la sencilla consiste en una búsqueda por forma tras de la que aparece una rueda de voces, ordenadas alfabéticamente, que pueden seleccionarse. El inicio de la consulta de *cap*, por ejemplo, ofrece las formas simples de *cap* como adv., m., m.f., prep., 1 indef., 2 indef., para pasar a las voces que incluyen la cadena fónica /cap/, como *capa* f., *capable* adj... hasta *capxirigany*; el usuario debe escoger. Es ésta otra forma de ofrecer el listado alfabético que se muestra de manera tradicional en los diccionarios con formato en papel, o a la izquierda de la pantalla, generalmente, cuando se trata de un cederrón.

La consulta compleja da como resultado una lista de entradas que cumple determinadas condiciones de selección. Por ejemplo, se puede preguntar por las palabras que comiencen por *no-* y que sean sustantivos femeninos, lo que proporciona los lemas no normativos **•no-agressió f.**, **•no-ingerència f.**, **•no-intervenció f.**, **•no-linealitat f.**, **•no-sinistralitat f.**, **no-violència f.** Frente a la lista de voces en orden alfabético que se ofrece al usuario, es difícil encontrar formas flexio-

guatge, Joan Costa, el envío de sus reflexiones en proceso de evaluación (“És actualment la «variació lingüística social» una categoria operativa per a la planificació lingüística de corpus?”) y del trabajo de JORDI BADÍA 2017.

nadas de alguno de sus lemas. La búsqueda de *nena*, femenino de *nen* (esp. *nene*, *nena*), ofrece la leyenda de que esa voz no está en el diccionario. Podría resolverse con facilidad esta prestación y convendría que se pusiera a punto, como otras posibilidades de hipertexto que aún no se ofrecen. Parece lógico que el hipertexto sea total en un diccionario digital, y la ventaja de este formato elegido es que pueden irse añadiendo nuevas prestaciones con el tiempo.

El hecho de no haberse elaborado este diccionario en orden alfabético, como era habitual antiguamente, sino por grupos afines de voces, y el extraordinario trabajo metódico previo que lo sustenta, así como la teoría lingüística y léxica que lo informa, le confieren un grado muy alto de homogeneidad y rigor notable en el tratamiento morfológico, sintáctico, semántico y temático de los artículos. Quien lo consulte se acostumbrará pronto a él, y lo que se encuentre no generará decepciones.

Lo que se ha venido señalando respecto de la información sintáctica y semántica hace patente que el diccionario no está dirigido a cualquier usuario. El catalán tiene un diccionario normativo para ello, el DIEC2, en última edición de 2007; el descriptivo está concebido para un público específicamente interesado en la lengua catalana: lingüistas, profesionales, investigadores, etc. El usuario, pues, debería contar con ciertos conocimientos previos sobre análisis de la lengua.

El DDLC aún no se ha presentado oficialmente, se encuentra, como se ha dicho, accesible totalmente en línea, pero aún no se ofrecen los textos de presentación y los créditos de los que han trabajado durante años en él; posiblemente la fecha de su presentación oficial tenga lugar muy pronto. Su director, Joaquim Rafel i Fontanals, y su coordinador, Joan Soler i Bou, han contado con un consejo de redacción compuesto de filólogos reconocidos, con un equipo de coordinación encargado de áreas específicas del proyecto y otro encargado de la redacción de artículos propiamente dicha. La nómina oficial de técnicos y redactores se compone de más de treinta lexicógrafos con otra larga lista de personas en formación que han trabajado en el diccionario en diferentes períodos de su gestación.

Por lo que se ha descrito y ejemplificado, el DDLC, además de ser un nuevo tipo de diccionario, es un gran diccionario; no en volumen ni en peso físico, sino por la cantidad de bytes e información ordenada, detallada y metódica que guarda. La descripción minuciosa del catalán contemporáneo permitirá el desarrollo de las labores normativas del Institut d'Estudis Catalans fundamentado en un conocimiento empírico de los usos lingüísticos. El catalán se dota así también de una descripción que sin duda será aprovechada para su procesamiento informático.

El DDLC es útil para todas las lenguas románicas y muy especialmente para la española, ya que cualquier trabajo léxico en esta lengua

se podría servir grandemente de los beneficios que aporta el diccionario, tal como hacen los trabajos gramaticales, morfológicos y sintácticos al emplear las gramáticas de las lenguas románicas. El diccionario también presenta, como se ha expuesto, novedades que son interesantes para la técnica lexicográfica en sí; es una apuesta novedosa y, lo que es bastante extraordinario, culminada.

REFERENCIAS

- BADIA I PUJOL, JORDI 2017. "Aproximació crítica a la nova gramàtica de l'Institut", *Els Marges*, 112, pp. 70-89.
- COSTA CARRERAS, JOAN en prensa. "La nova «tríada normativa» oficial de l'IEC: valoració i reflexions".
- DIEC2: INSTITUT D'ESTUDIS CATALANS 2007. *Diccionari de la llengua catalana*, Institut d'Estudis Catalans, Barcelona.
- GIEC: INSTITUT D'ESTUDIS CATALANS 2016. *Gramàtica de la llengua catalana*, Institut d'Estudis Catalans, Barcelona.
- INSTITUT D'ESTUDIS CATALANS. SECCIÓ FILOLÒGICA 2017. *Ortografia catalana*. Institut d'Estudis Catalans, Barcelona.