

El vocabulari especialitzat bàsic: el cas de les ciències naturals

ROSA ESTOPÀ
Universitat Pompeu Fabra
ORCID: 0000-0002-7382-1518
rosa.estopa@upf.edu

RESUM

Aquest article presenta el projecte «VEB. El vocabulari especialitzat bàsic en l'ensenyament obligatori; buidatge i anàlisi de la terminologia d'un corpus de llibres de textos escolars», que identifica i analitza la terminologia de ciències naturals dels llibres de text de l'ensenyament obligatori en català. L'objectiu del projecte VEB és detectar el lèxic especialitzat bàsic per millorar-ne la presència en diccionaris generals i escolars. El projecte ha creat el repositori digital VEB-WEB, que facilita la consulta terminològica per nivells educatius. Els resultats tenen aplicacions lexicogràfiques i didàctiques, i permeten proposar un diccionari escolar progressiu per millorar la comprensió lectora i l'aprenentatge científic.

PARAULES CLAU: vocabulari especialitzat bàsic; lexicografia especialitzada escolar; terminologia; diccionaris escolars



Rosa Estopà

és professora titular del Departament de Traducció i Ciències del Llenguatge i directora de la Càtedra Pompeu Fabra de la Universitat Pompeu Fabra. Catedràtica acreditada per l'Agència Nacional de Evaluación y Acreditación (ANECA), és doctora en Lingüística i investigadora al grup IULATERM (IULA-UPF), on coordina la línia de terminologia mèdica i comunicació en salut i codirigeix l'Observatori de Neologia. És vicepresidenta de la Societat Catalana de Terminologia i codirectora del màster en Terminologia. Actualment, és la investigadora principal del projecte de l'Institut d'Estudis Catalans Vocabulari Especialitzat Bàsic (ON-HEC).

TERMINÀLIA 32 (2025): 7-18 · DOI: 10.2436/20.2503.01.235
Data de recepció: 03/06/2025. Data d'acceptació: 14/10/2025
ISSN: 2013-6692 (impresa); 2013-6706 (electrònica) · <https://terminalia.iec.cat>

Abstract

Basic specialized vocabulary: The case of natural sciences

This article presents “VEB. Basic Specialized Vocabulary in Compulsory Education: A Corpus-Based Extraction and Analysis” project, which identifies and analyses science-related terminology from Catalan compulsory education textbooks. Its goal is to detect core-specialized vocabulary and assess its presence in general and school dictionaries. The project developed the open-access digital repository VEB-WEB, which allows users to consult terms by educational

level. The results have both lexicographic and educational applications, and support the creation of a progressive science dictionary to improve reading comprehension and scientific learning.

KEYWORDS: basic specialized vocabulary, school specialized lexicography, terminology, learner's dictionaries

1 Introducció

Des de finals del segle XIX, i coincidint amb l'impacte de la Revolució Industrial, les societats han viscut una transformació profunda que també ha afectat la relació amb el lèxic. En aquest context de canvi social accelerat —marcat per avenços tecnològics com la penicil·lina, el telègraf o la màquina de vapor— els parlants van començar a necessitar paraules específiques per descriure conceptes nous en àmbits tan diversos com la medicina, l'enginyeria o l'economia. Aquest fenomen no només s'ha mantingut, sinó que s'ha intensificat amb l'arribada de tecnologies modernes, com els ordinadors personals, la robòtica, Internet o, més recentment, la intel·ligència artificial.

Paral·lelament, el coneixement especialitzat ha deixat de ser un domini exclusiu d'una elit acadèmica o professional. D'una banda, gràcies a l'educació universal i, de l'altra, gràcies a la popularització de canals de divulgació —com programes científics de televisió, documentals o fòrums en línia—, termes abans restringits als especialistes han penetrat en el llenguatge comú. Així, unitats terminològiques com *emissió de carboni* o *algorisme*, que fins fa poc eren pròpies de científics i enginyers, ara es fan servir habitualment en discussions sobre canvi climàtic o xarxes socials. Per tant, sembla indiscutible l'existència d'un conjunt de paraules especialitzades fonamentals, conegudes per la població general, que podem denominar *vocabulari especialitzat bàsic* (VEB).

Aquest vocabulari és aquell que qualsevol parlant adult hauria de conèixer o podria necessitar conèixer. Tot i que no disposem d'una llista exacta dels termes que podem etiquetar com a fonamentals, és evident que aquest lèxic especialitzat ha anat creixent progressivament de manera exponencial. Aquest interès per l'especialització del coneixement es reflecteix, d'altra banda, en els diccionaris de llengua general, els quals incorporen un nombre cada vegada més gran d'entrades originalment no comunes, però que han arribat a difondre's socialment. Els lexicògrafs reconeixen que la tasca més

notable —i complexa— en l'actualització d'un diccionari consisteix en la incorporació de nous termes, tant en la selecció com en la representació (Estopà, 2022). Així, l'actualització constant de la terminologia respon a les necessitats canviants del parlant estàndard, com es fa evident en els pròlegs de la majoria dels diccionaris generals. Per exemple, el fragment següent del pròleg del *Grand Larousse* (Guilbert, Lagane, Niobey, 1989) reflecteix aquesta tendència, situant el nivell d'especialització com un element clau en el procés de selecció lèxica:

La total de termes dépasse largement celui de la partie lexicale du *Petit Larousse* qui est de 44.500. Le *Grand Larousse de la Langue Française* comprend, en plus, de **nombreux termes techniques**. (Gran Larousse, Guilbert, Lagane, Niobey, 1989)

Així doncs, partim de la premissa que la necessitat creixent de lèxic especialitzat en el parlant mitjà planteja reptes per al lexicògraf en relació amb l'objecte principal del diccionari de llengua. Un dels desafiaments més rellevants consisteix, doncs, a identificar quins són els termes fonamentals d'una llengua.

D'altra banda, el lèxic especialitzat que s'aprèn progressivament a l'escola constitueix l'embrió del coneixement científic i tècnic (Oxford University Press, 2018; Schuth et al., 2017). Sense la comprensió dels termes bàsics, els estudiants tenen serioses dificultats per entendre i assumir la informació que aquests termes vehiculen. Segons l'Informe PIRLS-2021 (publicat el 2023), Catalunya se situa quatre punts per sota de la mitjana espanyola en comprensió lectora, fet que evidencia la rellevància d'aquest problema. Si no s'entenen els termes, tampoc no s'entendrà el coneixement que aquests vehiculen ni se sabrà explicar.

La hipòtesi de treball planteja que el factor que iguala els membres d'una societat en relació amb la ciència i la tecnologia és el coneixement transmès a través de l'ensenyament obligatori, que arriba fins als setze anys. Durant aquest període educatiu, es transmet una quantitat considerable de coneixements científics i,

implícitament, molts termes que permeten referir-s'hi. Per aquest motiu, considerem que el vocabulari que es vehicula a través de l'ensenyament obligatori podria constituir la base per a la construcció del vocabulari especialitzat bàsic d'una llengua. En aquest sentit, el lèxic present en aquest nivell educatiu esdevé un recurs fonamental per a l'elaboració del que anomenem *vocabulari especialitzat bàsic* (VEB). D'altra banda, aquest conjunt lèxic hauria de ser considerat una terminologia imprescindible, tant per a la definició de la nomenclatura dels diccionaris de llengua general com per a la creació d'un diccionari d'aprenentatge de ciència. Finalment, cal destacar que un subconjunt d'aquest VEB és el que s'ha denominat *vocabulari acadèmic transversal*, és a dir, aquell format per termes abstractes i/o instrumentals que faciliten la comunicació sobre continguts de diverses disciplines (Schuth et al., 2017).

2 El projecte VEB (Vocabulari Especialitzat Bàsic)

2.1 Descripció

VEB és un projecte del **programa de recerca de l'Institut d'Estudis Catalans** (PRO2024-So4-ESTOPA) desenvolupat a l'Observatori de Neologia (grup IULATERM de la Universitat Pompeu Fabra), vehiculat a la Secció Filològica de l'IEC, durant el període 2023-2025. El projecte VEB neix justament d'aquestes constatacions i inquietuds i també del fet que, en llengua catalana, desconexim **quin és exactament el lèxic especialitzat que es transmet a l'escola obligatòria**. En altres contextos, com, per exemple, els Estats Units, fa anys que, en alguns estats, van establir un vocabulari especialitzat mínim en cada una de les àrees de coneixement obligatòries; es tracta d'uns cent termes bàsics per a cada un dels nivells educatius (Tennessee State Department of Education, 2006; Nagy i Herman, 1984; Marzano i Pickering, 2005; Carleton i Marzano, 2010). Per al català, tampoc no sabem si la terminologia bàsica que es vehicula a través de l'entorn escolar, sobretot a partir de currículums i llibres de text, es troba referenciada en el *Diccionari de la llengua catalana* de l'IEC i en altres *diccionaris d'aprenentatge* de referència destinats als escolars.

Davant d'aquestes incerteses, el projecte VEB planteja diverses preguntes de recerca, entre les quals destaquem:

1. Per què un terme és al diccionari general mentre que un altre terme no hi és?
2. Quins criteris objectius té el lexicògraf per incorporar terminologia en un diccionari de llengua general?
3. La terminologia que es vehicula a través de l'ensenyament obligatori està representada en els diccionaris de llengua general? I en els escolars?

4. Quina és la terminologia que es vehicula a través de l'ensenyament obligatori?
5. Quin grau de progressió i de repetició lèxica hi ha en els diferents cursos i cicles de l'ensenyament obligatori?

2.2 Objectius

Els objectius generals del projecte VEB són:

- Identificar el vocabulari especialitzat bàsic utilitzat en els llibres de text de l'ensenyament obligatori en català, seqüenciat per cursos i per àrees de coneixement.
- Analitzar la representació del VEB en els diccionaris de referència per a la llengua catalana, determinant quins termes no estan recollits en el DIEC2 o en diccionaris escolars. En conseqüència, aquests termes poden considerar-se lexicogràficament neològics des de la perspectiva de l'usuari general, d'acord amb els postulats teòrics de Cabré (2015), Guerrero (2017) i Estopà (2021).

2.3 Finalitats

El projecte VEB té dues finalitats clarament diferenciades:

- Finalitat lexicogràfica:
 - Els resultats proporcionen dades objectives que faciliten la inclusió sistemàtica de terminologia en els diccionaris de llengua general.
 - Aquests resultats són, a més, valuosos per a l'elaboració d'un diccionari de ciència escolar o bàsic a partir de necessitats d'ús real.
- Finalitat didàctica:
 - Les dades permeten valorar la idoneïtat de la terminologia emprada en els llibres de text en cada nivell educatiu.
 - També estableixen criteris per establir la terminologia bàsica, mínima i adequada per a cada nivell escolar.

2.4 Metodologia

Durant els anys 2023 i 2024, el projecte VEB s'ha centrat en l'estudi de la terminologia de l'àrea de coneixement de les **ciències naturals** (Estopà, 2024), mitjançant l'anàlisi de la terminologia d'un corpus de llibres de text utilitzats en els cursos obligatoris de l'ensenyament primari i secundari. A partir del 2025 s'ha començat l'estudi de la terminologia de l'àrea de coneixement de les matemàtiques.

Tot i que molts centres educatius han reduït l'ús de llibres de text a les aules, encara n'hi ha un nombre significatiu que en continuen fent ús. A més, aquests materials segueixen sent una referència fonamental per

a molts professionals del sector educatiu. En aquest sentit, els llibres de text ofereixen una base objectiva i estructurada que unifica els continguts per a cada nivell escolar, motiu pel qual el projecte ha optat per partir d'aquesta font comuna i objectivable per tal d'estructurar i analitzar la terminologia de manera sistemàtica.

Les fases de treball del projecte han estat les següents:

1. Selecció dels llibres de text representatius i dels currículums per nivells escolars i per matèries.
2. Constitució d'un corpus textual seqüenciat per nivells escolars i per matèries.
3. Buidatge de la terminologia de cada subcorpus amb informació de la freqüència d'ús.
4. Anàlisi i classificació de la terminologia segons la presència en diferents obres lexicogràfiques generals i especialitzades de referència.

Fases 1 i 2: selecció i constitució del corpus textual

El projecte, amb l'assessorament d'experts en educació, va dur a terme la selecció dels llibres de referència més utilitzats a les escoles catalanes per a les assignatures de ciències del medi, en l'educació primària, i de biologia i geologia, així com física i química, en l'educació secundària obligatòria (ESO). Els llibres seleccionats compren publicacions editades entre els anys 2015 i 2025.

Concretament, es van analitzar 48 llibres de primària procedents de 8 editorials: Barcanova, Baula, Castellnou, Cruïlla, Edebé, La Galera, Santillana i Vicens Vives. Pel que fa a l'ESO, es van estudiar 28 llibres de 7 editorials: Barcanova, Casals, Cruïlla, McGraw-Hill, Santillana, Teide i Vicens Vives. En total, per a aquesta àrea de coneixement, s'ha realitzat l'anàlisi terminològica de 76 llibres.

Fase 3: buidatge de la terminologia

Es va optar per realitzar el buidatge de la terminologia dels llibres de text de manera manual. Aquesta decisió es va prendre perquè els llibres de text inclouen nombroses taules, esquemes i il·lustracions amb una gran quantitat de terminologia, fet que complicava significativament el tractament automatitzat mitjançant extractors de terminologia, especialment en el cas dels llibres de primària.

Cada terme només s'ha recollit una sola vegada per llibre i s'ha creat una fitxa per a cada ocurrència diferent que apareixia en un llibre de text. Això permet identificar en quins textos i en quins cursos es troba un terme determinat.

Fase 4: anàlisi de la terminologia

Per a cada terme diferent se n'ha registrat la categoria gramatical, l'àrea temàtica i si estava recollit d'alguna manera en obres lexicogràfiques de referència. Concretament, s'ha verificat si el terme estava recollit en el *Diccionari de la llengua catalana* de l'Institut d'Estudis Catalans (DIEC2) o en el *DIDAC. Diccionari escolar* (DIDAC) d'Enciclopèdia Catalana.

Malgrat que tota la terminologia analitzada és usada en els llibres de text de ciències naturals, es va constatar que, a més dels termes de temes de biologia, química o física, també hi ha una gran quantitat de terminologia més pròpia d'altres àrees de coneixement. Algunes d'aquestes són àrees afins com ara la medicina, la geologia, l'astronomia o la meteorologia, mentre que d'altres són més allunyades, com el dret, la religió o la lingüística. A més, es va identificar la presència de termes transversals a moltes àrees del coneixement, així com termes instrumentals característics del lèxic de la ciència i dels mètodes científics. Per aquest motiu, es va considerar oportú establir una bateria de marques temàtiques basades en les utilitzades per altres diccionaris, especialment, pel DIEC2 (Estopà, 1998).

Pel que fa a la presència d'un terme en els diccionaris, es distingeixen les possibilitats següents:

- **Sí:** quan el terme apareix com a entrada o subentrada.
- **No:** quan el terme no es troba en el diccionari.
- **Exemple:** quan el terme apareix només com a exemple d'una accepció.
- **Definició:** quan el terme forma part de la definició d'una paraula afí.
- **Làmina:** quan el terme apareix com a part d'una il·lustració.

Per tal de gestionar les dades de manera més sistemàtica i objectiva, s'ha desenvolupat el programa Gestor-VEB, una eina que facilita el treball dels investigadors. Aquesta plataforma preveu dos tipus d'usuari: administrador i investigador. L'usuari administrador té capacitat per modelar tots els camps previstos, mentre que l'usuari investigador pot introduir termes nous (o ocurrències d'un terme existent a la base de dades), modificar-los i suprimir-los. Per garantir la sistematització de les dades i reduir els errors humans, el gestor funciona amb menús desplegable. El programa també inclou un mòdul de consulta avançada que permet recuperar informació mitjançant el creuament de tots els camps disponibles. A més, s'ha previst la possibilitat d'importar dades massives des d'un arxiu Excel i d'exportar dades en diversos formats. La figura 1 mostra la pantalla principal del Gestor-VEB:

Termes

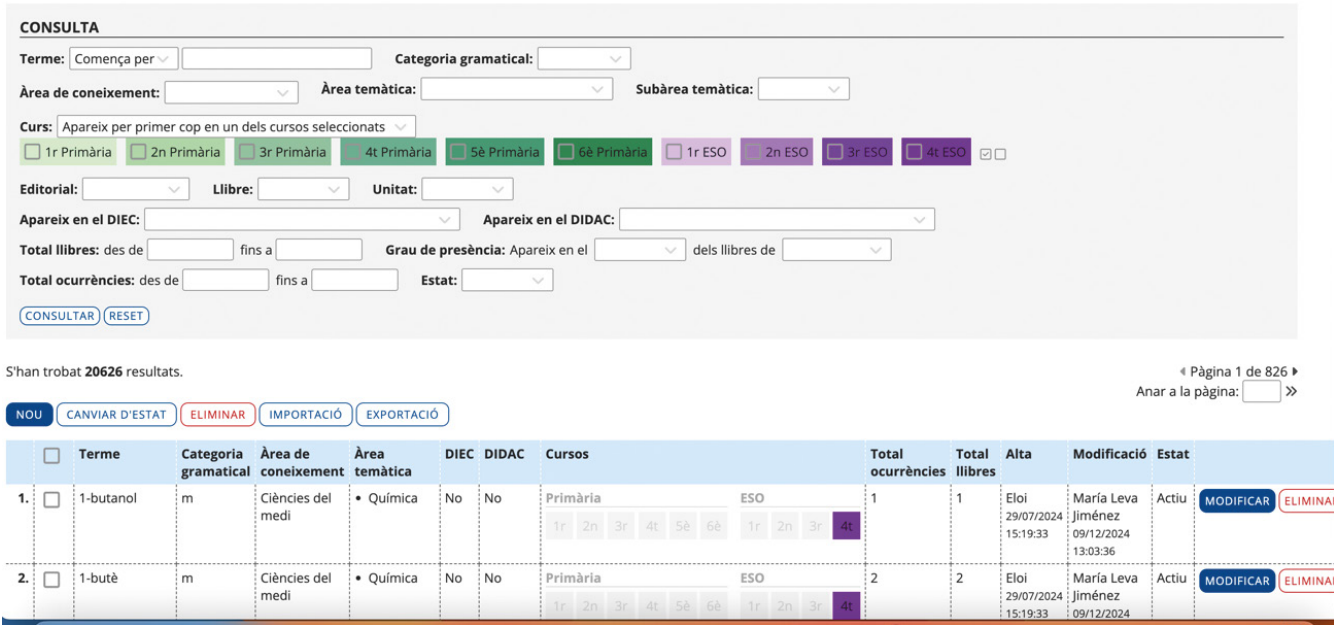


FIGURA 1. Pantalla principal del Gestor-VEB

3 El VEB-WEB, repositori en línia

Les dades recollides i analitzades es posen a disposició mitjançant el portal en línia VEB-WEB, d'accés obert, que permet a qualsevol persona interessada —especialment lexicògrafs, terminòlegs o especialistes en educació— consultar de manera amigable els resultats del projecte. Aquest repositori recopila el **vocabulari especialitzat bàsic** que es fa servir en l'ensenyament obligatori a Catalunya, des del primer curs de l'educació primària fins al quart curs de l'educació secundària obligatòria (ESO).

El repositori inclou aquesta terminologia, classificada per nivells educatius, extreta d'una selecció representativa de llibres de text i materials didàctics en català de diverses matèries obligatòries. A més, les dades del projecte VEB s'actualitzen de manera contínua per garantir-ne la pertinència i la utilitat. L'adreça electrònica del portal és: <https://veb.iula.upf.edu/>.

El VEB-WEB s'organitza en una primera pantalla que ofereix: a) dades generals dels resultats del projecte, b) un cercador bàsic, c) un cercador avançat, i d) un mòdul de dades estadístiques.

a) Dades generals

A principis del 2025, el VEB-WEB conté 20.626 termes diferents de l'àrea de coneixement de ciències naturals (durant el 2025 s'està treballant en l'àrea de coneixement de matemàtiques). Actualment, en el segon semestre de 2025, el repositori té 26.749 termes diferents, corresponents a 124 llibres de text, de dues àrees de coneixement, de 15 editorials. Les dades s'actualitzen permanentment.

A la primera pantalla del portal s'ofereixen algunes dades estadístiques (figura 2). Per exemple, els percentatge dels termes que apareixen per primer cop en un nivell educatiu (els gràfics corresponen tots al repositori durant el mes de gener de 2025):

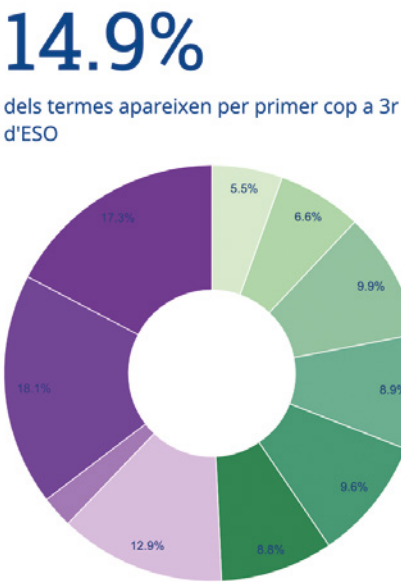


FIGURA 2. Percentatge del termes nous en cada curs educatiu

També s'hi poden consultar els termes que apareixen a tots els nivells educatius de primària i els que surten a tots els nivells educatius de secundària (figura 3):

abella f

i 307 termes més apareixen en tots els cursos de Primària

abella f, acer m, adult adulta m i f, aeroport m, aigua f, aigua dolça f, aigua líquida f, aigua potable f, aire m, ajuntament m, ala f, alçada f, aleta f, alga f, aliment m, alimentació f, alzina f, amfibi m, animal m, animal herbívor m, animal invertebrat m, animal vertebrat m, any m, aparell m, aparell elèctric m, arbre m, arbre frutífer m, ...

[Veure la llista completa](#)

abella f

i 1103 termes més apareixen en tots els cursos d'ESO

abella f, abocador m, absorció f, acció f, acer m, àcid m, àcid desoxiribonucleic m, àcid nucleic m, àcid oleic m, àcid ribonucleic m, acidesa f, activitat volcànica f, adaptació f, ADN m, adob m, adrenalina f, aerogenerador m, agricultor m, agricultora m i f, agricultura f, àguila f, aigua f, aigua de mar f, aigua destil·lada f, aigua dolça f, aigua líquida f, aigua marina f, aigua mineral f, ...

[Veure la llista completa](#)

FIGURA 3. Termes que apareixen a tots els cursos de primària i a tots els d'ESO

b) Cercador bàsic

El cercador simple, que es localitza en la part superior de la pàgina inicial, permet cercar termes de manera exacta (que sigui igual), per aproximació (que comen-

ci, que acabi, que contingui) o mitjançant la llista dels termes ordenats alfabèticament, tal com es mostra en la figura 4.

c) Cercador avançat

El cercador avançat permet fer cerques més complexes creuant diversos camps: categoria gramatical, àrea de coneixement, nivell educatiu, aparició en el DIEC o en el DIDAC, àrea temàtica i percentatge de presència en els llibres d'un nivell educatiu determinat. Per exemple, en el cas del nivell educatiu, es pot seleccionar una opció entre el menú següent:

- Apareix per primer cop en un dels cursos seleccionats.
- Es fa servir en algun dels cursos seleccionats.
- Es fa servir com a mínim en tots els cursos seleccionats.
- Es fa servir només en els cursos seleccionats.
- No es fa servir en cap dels cursos seleccionats.

I en el cas de l'àrea temàtica, es desplega un menú amb totes les àrees temàtiques incloses en el projecte. La captura de pantalla de la figura 5 mostra la pàgina principal del cercador avançat.

Una vegada es fa una consulta, els resultats es presenten amb unes capses virtuals amb les dades més rellevants: la categoria gramatical, quina és la seva presència en les dues obres de referència seleccionades i en quins nivells educatius s'ha registrat el terme. Ara bé, quan s'entra en qualsevol terme, la informació augmenta i s'ofereixen les àrees temàtiques associades, la representació en els diversos nivells educatius i les fonts de referència específiques de cada curs (vegeu la figura 6).



FIGURA 4. Pantalla amb l'accés al cercador bàsic

CERCADOR

Terme: Comença per Categoria gramatical: Àrea de coneixement:

Curs: Apareix per primer cop en un dels cursos seleccionats Primària: ☐ 1r ☐ 2n ☐ 3r ☐ 4t ☐ 5è ☐ 6è ESO: ☐ 1r ☐ 2n ☐ 3r ☐ 4t ☐

Apareix en el DIEC: Apareix en el DIDAC:

Àrea temàtica: Grau de presència: Apareix en el dels llibres de

FIGURA 5. Pantalla principal del cercador avançat

Per exemple, si consultem quins termes contenen la paraula *aigua*, el sistema ens retorna que hi ha 93 termes que contenen la seqüència *aigua* i ens les ofereix en capsetes virtuals en diverses pantalles. Reproduïm a la figura 7 la primera pantalla dels resultats que es mostren en 24 capsetes a cada pantalla.

Aquestes 93 unitats terminològiques són: *aigua*, *aigua atmosfèrica*, *aigua blana*, *aigua bruta*, *aigua bullent*,

aigua calenta, *aigua càlida*, *aigua carbonatada*, *aigua carbònica*, *aigua congelada*, *aigua consumida*, *aigua contaminada*, *aigua continental*, *aigua continental corrent*, *aigua continental dolça*, *aigua continental estancada*, *aigua corrent*, *aigua d'alta salinitat*, *aigua d'escolament*, *aigua d'escorriment*, *aigua de baixa salinitat*, *aigua de brom*, *aigua de clor*, *aigua de colònia*, *aigua de consum*, *aigua de deu*, *aigua de iode*, *aigua de l'aixeta*, *aigua de la pluja*, *aigua de mar*, *aigua de pluja*,

AIGUA CORRENT f

TORNAR

Diccionaris



Àrees temàtiques



Cursos Primària

1r 0 llibres	2n 0 llibres	3r 2 llibres	4t 1 llibre	5è 1 llibre	6è 1 llibre
-----------------	-----------------	-----------------	----------------	----------------	----------------

Cursos ESO

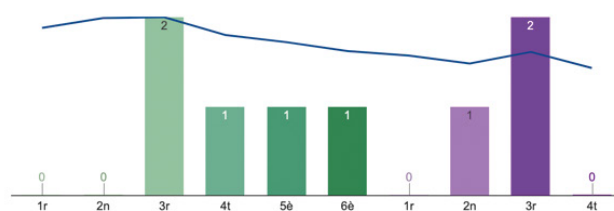
1r 0 llibres	2n 1 llibre	3r 2 llibres	4t 0 llibres
-----------------	----------------	-----------------	-----------------

Llista d'ocurrències

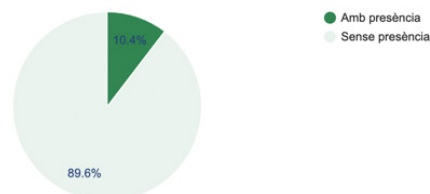
3r Primària
• Coneixement del medi (Editorial Santillana - ISBN 978-84-7918-762-0)
• Coneixement del medi natural (Editorial Barcanova - ISBN 978-84-489-5834-3)
4t Primària

Estadístiques

Ocurrències del terme per curs i mitjana d'ocurrències per curs



Primària > Presència als llibres



ESO > Presència als llibres

FIGURA 6. Informació del registre *aigua corrent*

S'han trobat 93 resultats.

« Pàgina 1 de 4 »
Anar a la pàgina: »

aigua f DIEC: Sí · DIDAC: Sí Primària: 1r 2n 3r 4t 5è 6è ESO: 1r 2n 3r 4t	aigua carbònica f DIEC: Exemple · DIDAC: No Primària: 1r 2n 3r 4t 5è 6è ESO: 1r 2n 3r 4t	aigua corrent f DIEC: Sí · DIDAC: No Primària: 1r 2n 3r 4t 5è 6è ESO: 1r 2n 3r 4t
aigua atmosfèrica f DIEC: No · DIDAC: No Primària: 1r 2n 3r 4t 5è 6è ESO: 1r 2n 3r 4t	aigua congelada f DIEC: No · DIDAC: No Primària: 1r 2n 3r 4t 5è 6è ESO: 1r 2n 3r 4t	aigua d'alta salinitat f DIEC: No · DIDAC: No Primària: 1r 2n 3r 4t 5è 6è ESO: 1r 2n 3r 4t
aigua blana f DIEC: Sí · DIDAC: No Primària: 1r 2n 3r 4t 5è 6è ESO: 1r 2n 3r 4t	aigua consumida f DIEC: No · DIDAC: No Primària: 1r 2n 3r 4t 5è 6è ESO: 1r 2n 3r 4t	aigua d'escolament f DIEC: No · DIDAC: No Primària: 1r 2n 3r 4t 5è 6è ESO: 1r 2n 3r 4t
aigua bruta f DIEC: No · DIDAC: No Primària: 1r 2n 3r 4t 5è 6è ESO: 1r 2n 3r 4t	aigua contaminada f DIEC: No · DIDAC: No Primària: 1r 2n 3r 4t 5è 6è ESO: 1r 2n 3r 4t	aigua d'escorriment f DIEC: No · DIDAC: No Primària: 1r 2n 3r 4t 5è 6è ESO: 1r 2n 3r 4t

FIGURA 7. Resultats de cercar els termes que comencen per *aigua*

aigua de refrigeració, aigua de reg, aigua de reg agrícola, aigua de riu, aigua del mar, aigua depurada, aigua desionitzada, aigua destil·lada, aigua dolça, aigua domèstica, aigua dura, aigua embassada, aigua embotellada, aigua estancada, aigua eutròfica, aigua evaporada, aigua fluorada, aigua freàtica, aigua freda, aigua gasosa, aigua gelada, aigua glaçada, aigua hidrogràfica, aigua líquida, aigua marina, aigua mineral, aigua mineral amb gas, aigua mineral embotelladora, aigua mineral sense gas, aigua neta, aigua oceànica, aigua oxigenada, aigua potable, aigua pura, aigua quieta, aigua ramadera, aigua refrigerada, aigua residual, aigua residual urbana, aigua sabonosa, aigua salabrosa, aigua salada, aigua salada del mar, aigua salina, aigua salobre, aigua salvatge, aigua sanitària, aigua sòlida, aigua subterrània, aigua superficial, aigua tèbia, aigua termal, aigua urbana, aigua-ros, aiguader, aiguadera, aiguafort, aiguamarina, aigumoll, aiguanex, aigurràs i aiguat.

Podem refinar la cerca restringint qualsevol dels camps que s'hi inclouen. Per exemple, podríem demanar que només ens donés aquells termes que contenen el mot aigua i que surten a tots els cursos educatius i el resultat seria que només tres termes compleixen aquesta premissa: aigua, aigua dolça i aigua líquida. O només aquelles unitats lèxiques amb aigua que apareixen a tota la primària i serien els tres termes anteriors

més aigua potable. O tots els termes amb aigua que trobem a tots els quatre cursos de secundària i augmentaríem a 9 unitats, tal com es mostra en la figura 8.

O, per exemple, que el filtre fos només aquelles unitats que contenen aigua i que no estan recollides en el DIEC i els resultats són que, de les 93 unitats candidates, 58 (62 %) no estan recollides en el DIEC2 (que hem marcat amb negreta), ni com a entrada ni com a subentrada, exemple o definició: aigua, **aigua atmosfèrica**, aigua blana, **aigua bruta**, aigua bullent, **aigua calenta**, **aigua càlida**, **aigua carbonatada**, aigua carbònica, **aigua congelada**, **aigua consumida**, **aigua contaminada**, **aigua continental**, **aigua continental corrent**, **aigua continental dolça**, **aigua continental estancada**, aigua corrent, **aigua d'alta salinitat**, **aigua d'escolament**, **aigua d'escoriment**, **aigua de baixa salinitat**, **aigua de brom**, aigua de clor, aigua de colònia, **aigua de consum**, **aigua de deu**, **aigua de iode**, **aigua de l'aixeta**, aigua de la pluja, aigua de mar, aigua de pluja, **aigua de refrigeració**, **aigua de reg**, **aigua de reg agrícola**, aigua de riu, aigua del mar, **aigua depurada**, aigua desionitzada, aigua destil·lada, aigua dolça, **aigua domèstica**, aigua dura, **aigua embassada**, **aigua embotellada**, **aigua estancada**, **aigua eutròfica**, **aigua evaporada**, **aigua fluorada**, aigua freàtica, **aigua freda**, **aigua gasosa**, **aigua gelada**, **aigua glaçada**, **aigua hidrogràfica**, **aigua líquida**,

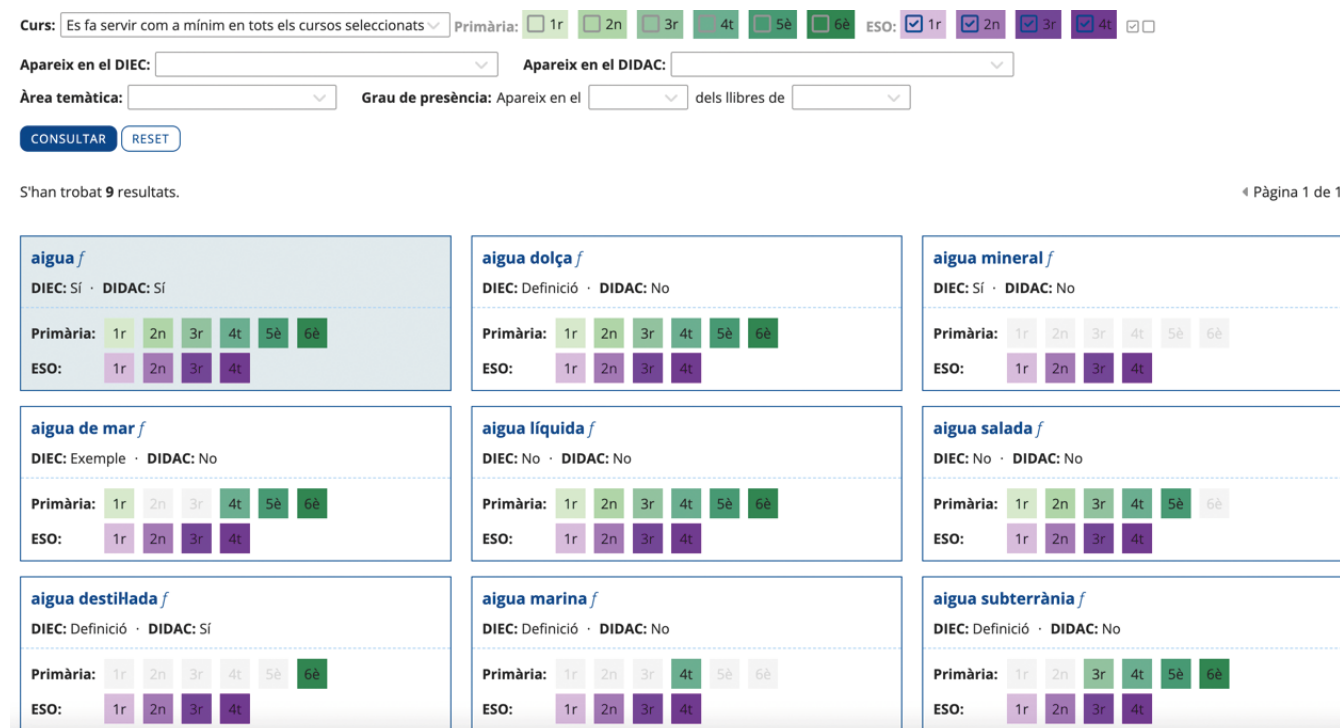


FIGURA 8. Resultats de cercar els termes que contenen aigua a tots els cursos de l'ESO

aigua marina, aigua mineral, **aigua mineral amb gas**, **aigua mineral embotelladora**, **aigua mineral sense gas**, **aigua neta**, **aigua oceànica**, aigua oxigenada, aigua potable, aigua pura, **aigua quieta**, **aigua ramadera**, **aigua refrigerada**, aigua residual, **aigua residual urbana**, **aigua sabonosa**, **aigua salabrosa**, **aigua salada**, **aigua salada del mar**, aigua salina, aigua salobre, **aigua salvatge**, **aigua sanitària**, **aigua sòlida**, aigua subterrània, aigua superficial, **aigua tèbia**, **aigua termal**, **aigua urbana**, aigua-ros, aiguader, **aiguadera**, **aiguafort**, **aiguamarina**, **aiguamoll**, **aiguaneix**, **aiguarràs** i **aiguat**.

Si filtrem per quines són al DIDAC, trobem que només 11 (12 %) d'aquestes unitats terminològiques es recullen en aquest diccionari escolar. I podríem fer molts altres filtres segons els camps que restringís i creuéssim.

d) Mòdul d'estadístiques

Finalment, el VEB-WEB inclou un mòdul d'estadístiques que ofereix una sèrie de dades representades gràficament que poden ser útils per a l'usuari. En primer lloc, podem saber la quantitat de termes usats en cada curs contrastat amb els termes nous respecte del curs anterior, tal com es mostra en la figura 9.

Termes utilitzats en cada curs i termes nous respecte del curs anterior

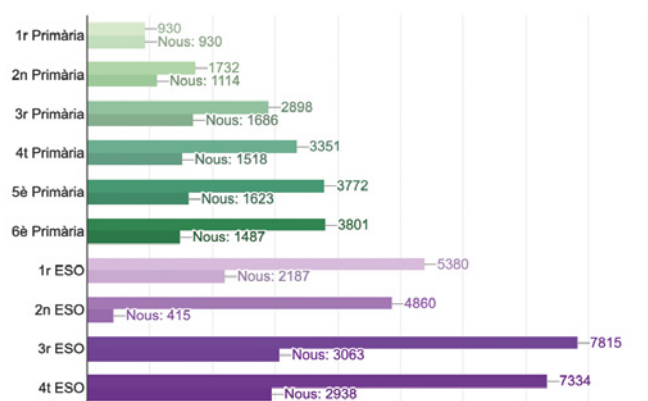


FIGURA 9. Termes utilitzats en cada curs i termes nous respecte del curs anterior

La informació que recull el gràfic anterior permet conèixer l'augment progressiu de cada nivell cognitiu i veure que a primer, a segon i a tercer de primària el

lèxic nou suposa el 64 %, el 62 % i el 58 %, respectivament, mentre que a la resta dels cursos se situa en una forquilla d'entre el 39 % i el 45 % (4t-P: 45 %, 5è-P: 43 %, 6è-P: 39 %, 1r-ESO: 40 %, 3r-ESO: 39 % i 4t-ESO: 40 %). Hi ha una excepció que és segon de l'ESO, en què el lèxic nou respecte del curs anterior és només del 8,5 %. Caldria investigar aquesta dada sorprenent perquè no tenim un argument coherent.

També podem saber gràficament com es distribueix el vocabulari de cada curs i el percentatge de termes d'un curs que ja s'han vist per primera vegada en un altre. Així es pot conèixer, per exemple, a primer de l'ESO quin és el percentatge que ja vam veure a primer, quin ja vam veure a segon, a tercer, a quart i a cinquè de primària, de manera que constatem quin és el grau de progressió i de repetició terminològica en cada nivell educatiu, tal com es mostra al gràfic següent de la figura 10:

Curs de la primera aparició dels termes utilitzats

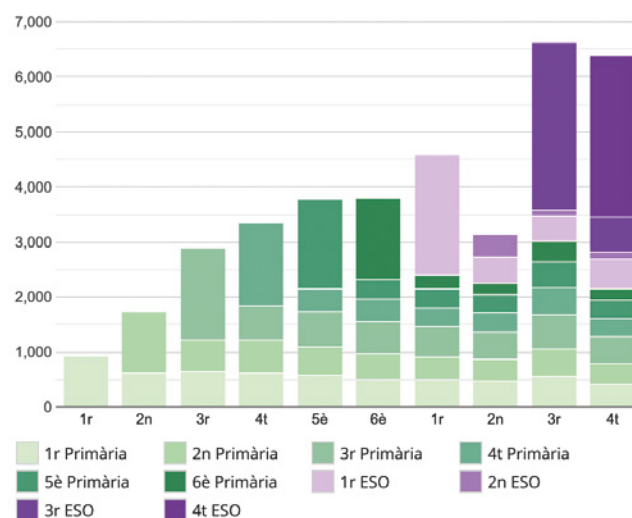


FIGURA 10. Curs de la primera aparició des termes per cada nivell educatiu

Així, també tenim estadístiques que tenen a veure amb la representativitat dels termes en el diccionari (vegeu la figura 11).

Aparició en els diccionaris dels termes utilitzats en cada curs

■ DIEC + DIDAC ■ DIEC ■ DIDAC ■ Cap diccionari ■ Informació incompleta

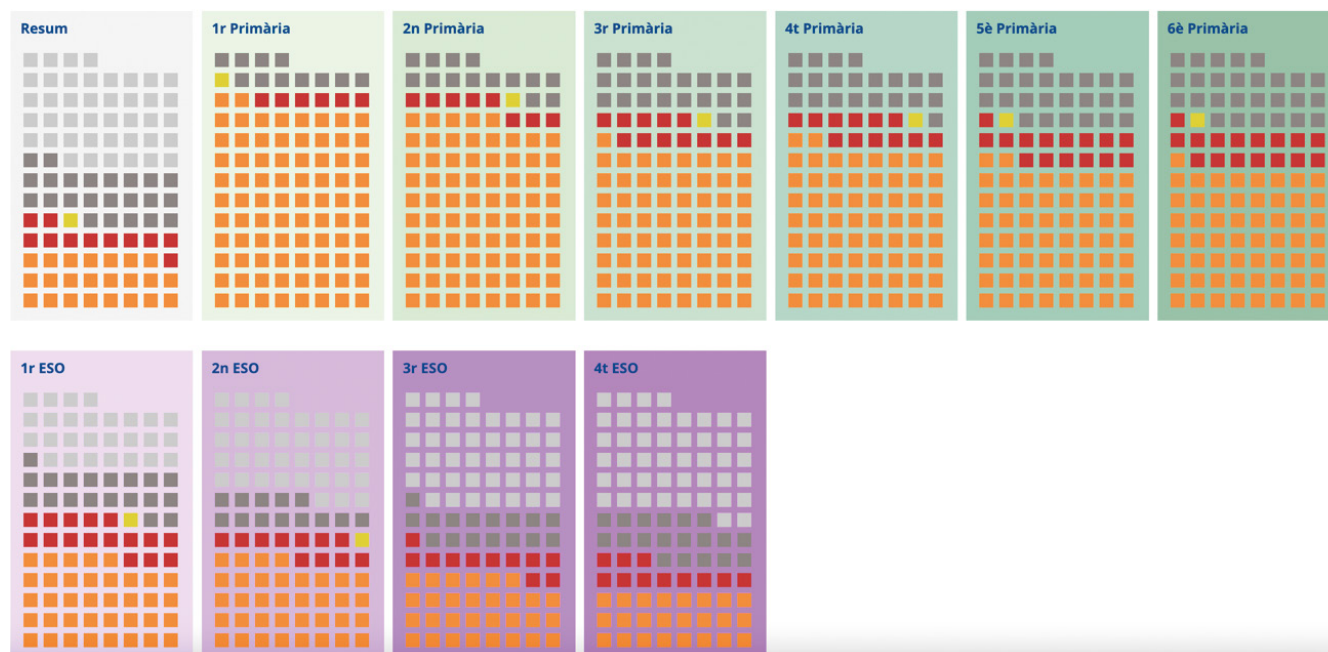


FIGURA 11. Pantalla principal de l'aparició en els diccionaris dels termes per nivells educatius

En quart lloc, podem saber el percentatge de llibres buidats per a cada curs en una àrea de coneixement determinat, en aquest cas en ciències del medi (figura 12):

Nombre de llibres utilitzats per curs amb termes de Ciències del medi

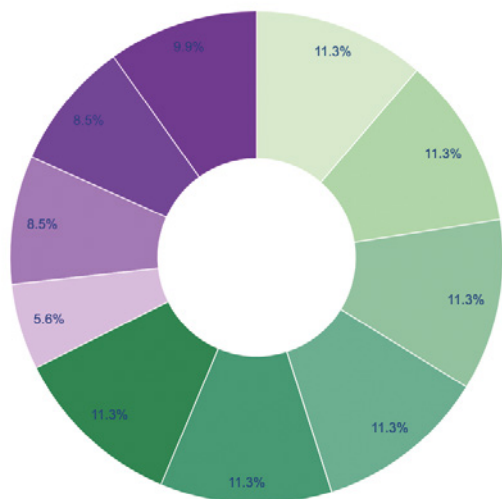


FIGURA 12. Nombre de llibres estudiats per cada curs de l'àrea de ciències del medi

Finalment, també podem veure en un gràfic el percentatge de termes en cada una de les àrees del co-

neixement. En aquest sentit, sabem que les sis àrees temàtiques més freqüents dels termes que hem recollit en els llibres de ciències del medi s'associen, per ordre, a la medicina, la química, la biologia, la física, la zoologia i la geologia.

4 Observacions finals

Els resultats del projecte VEB indiquen que la terminologia recollida en els llibres de text proporciona dades rellevants sobre candidats a termes bàsics, els quals podrien integrar tant el diccionari de llengua general com ser la nomenclatura objectiva d'un diccionari bàsic de ciència. Aquestes dades, no obstant això, susciten un debat sobre la conveniència d'incloure les unitats polilèxiques com a entrades independents, especialment en el context dels nous formats lexicogràfics. Per exemple, si bé s'ha observat una notable presència d'unitats polilèxiques amb el nucli *aigua*, cap no té cabuda com a entrada independent en els diccionaris generals actuals i moltes d'aquestes unitats tampoc no la tenen com a subentrada. La raó principal era l'espai. Aquest criteri no té ja sentit en la lexicografia del segle XXI, en què els diccionaris són, majoritàriament, digitals. El mateix passa amb altres unitats com les sigles o certs noms propis, com, per exemple, el nom dels planetes.

Les dades obtingudes resulten, a més, molt útils per revisar sistemàticament les subentrades i els exemples dels termes fonamentals. Això permetria controlar les sèries d'unitats polilèxiques que comparteixen el mateix hiperònim i equilibrar els exemples. També facilitaria la gestió de la variació de les unitats terminològiques i de les variants més utilitzades. Tenir la terminologia que es vehicula en els llibres de text analitzada lexicogràficament resulta també molt útil per revisar sistemàticament les subentrades i els exemples dels termes fonamentals, cosa que facilita tant el control de les sèries d'unitats polilèxiques com la gestió de la variació terminològica i de les variants més utilitzades.

Així mateix, aquests resultats obren la porta a reflexions didàctiques importants, ja que permeten extreure llistes de termes que podrien establir els mínims imprescindibles de conceptes per a cada nivell escolar, cosa que contribuiria a una introducció progressiva i

equilibrada del coneixement. D'altra banda, el fet que una part significativa del lèxic present als llibres de text no es trobi en els diccionaris generals posa de manifest la necessitat de desenvolupar diccionaris especialitzats de ciència destinats a l'aprenentatge. En aquest context, les dades obtingudes podrien ser fonamentals per construir diccionaris progressius estructurats en tres nivells:

- **Elemental:** de primer a tercer de primària.
- **Bàsic:** de quart a sisè de primària.
- **Essencial:** de primer a quart d'ESO.

En definitiva, comptar amb una terminologia extreta de llibres de text de diversos nivells constitueix una base de gran valor per al desenvolupament de recursos lexicogràfics i per millorar l'ensenyament sobretot pel que fa a la comprensió lectora i l'aprenentatge científic dels conceptes fonamentals. ✿

Bibliografia

- CABRÉ, M. Teresa (2015). «La neologia: un nou camp a la cerca de la seva consolidació científica». *Caplletra*, núm. 59, p. 125-136.
- CARLETON, Lindsay; MARZANO, Robert (2010). *Vocabulary games for the classroom*. Bloomington: Marzano Research Laboratory.
- DIEC2 = INSTITUT D'ESTUDIS CATALANS (2007). *Diccionari de la llengua catalana* [en línia]. 2a ed. Barcelona: Edicions 62: Enciclopèdia Catalana. <<https://dlc.iec.cat/>>.
- ESTOPÀ, Rosa (1998). «El léxico especializado en los diccionarios de lengua general: marcas temáticas». *Revista Española de Lingüística*, vol. 28, núm. 2, p. 359-387.
- (2021). *El diccionario escolar de ciencia*. Sevilla: McGraw Hill, 2021.
- (2022). «The general language dictionary in light of advances: Specialised neologisms». A: FREIXA, Judit [et. al.] (ed). *Dictionarization of Catalan neologisms*. Nova York: Peter Lang, p. 295-312. (Catalan Studies in Culture and Linguistics; 3)
- (2024). «Neología de receptor: estudio progresivo de la terminología que se vehicula a través de los libros de texto de la enseñanza obligatoria». A: ROSSI, M. (ed.). *Néologie, terminologie et variation = Neología, terminología y variación = Neologia, terminologia e variazione*. Lausana: Peter Lang, p. 107-122.
- GRUP ENCICLOPÈDIA (2008). *DIDAC: Diccionari escolar* [en línia]. Barcelona: Enciclopèdia Catalana. <<https://www.diccionari.cat/didac>>.
- GUERRERO, Gloria (2017). «Nueva orientación de la terminología y la neología en el ámbito de la semántica léxica». *RILCE*, vol. 33, núm. 3, p. 1385-1415.
- Le Grand Larousse*. París: Larousse, 1989.
- MARZANO, Robert J.; PICKERING, Debra (2005). *Building academic vocabulary: Teacher's manual*. Virginia: Association for Supervision and Curriculum Development.
- MINISTERIO DE EDUCACIÓN Y FORMACIÓN PROFESIONAL (2023). *PIRLS 2021 Estudio internacional de progreso en comprensión lectora* [en línia]. S. Il.: Secretaría General Técnica; Subdirección General de Atención al Ciudadano, Documentación y Publicaciones <https://www.libreria.educacion.gob.es/ebook/175187/free_download/>
- NAGY, Williams E.; HERMAN, Patricia (1984). *Limitations of vocabulary instruction*. Urbana: University of Illinois.
- OXFORD UNIVERSITY PRESS (2018). *Why closing the word gap matters: Oxford language report*. Oxford: Oxford University Press.

SCHUTH, Elisabeth; KÖHNE, Judith; WEINERT, Sabine (2017). «The influence of academic vocabulary knowledge on school performance». *Learning and Instruction*, vol. 49, p. 157-165.

TENNESSEE DEPARTMENT OF EDUCATION (2006). «Tennessee academic vocabulary: A guide for Tennessee educators». <<http://dr-hatfield.com/EDUC536/docs/VOCABULARY.pdf>>. [Consulta: 1 octubre 2006].

Nota

1. **VEB** és un projecte del **programa de recerca de l'Institut d'Estudis Catalans** (PRO2024-So4-ESTOPA) desenvolupat a l'Observatori de Neologia (grup IULATERM de la Universitat Pompeu Fabra) durant el període 2023-2025. Direcció: Rosa Estopà. Ponent de l'IEC: Mercè Lorente. Equip de treball: Mariona Arnau-Garcia, Elisenda Bernal, Pedro J. Bueno, Ona Domènech-Bagaria, Rosa Estopà, Judit Freixa, Elisabet Llopart-Saumell, Ludovic Masson, Amor Montané, Albert Morales, Jorge M. Porras-Garzón, Laia Vidal-Sabanés. Personal de suport a la recerca (PSR) 2024: Eloi Singla i Milian. Pàgina web: <https://www.upf.edu/web/vocabulari-especialitzat-basic>.