

1. Introducció

L'elaboració d'un **nou protocol** de buidatge s'ha fet necessària perquè amb els anys han anat canviant les eines de treball i altres aspectes de la feina dels observatoris de neologia. Aquest document reuneix les indicacions perquè qualsevol col·laborador de qualsevol node de les xarxes d'observatoris pugui fer buidatge. El símbol ● indica que hi ha un canvi important respecte del protocol anterior.

Fer **buidatge neològic** és extreure d'un text oral o escrit les unitats lèxiques considerades neològiques seguint el criteri lexicogràfic, és a dir, considerant neològiques les paraules que no apareixen consignades de manera idèntica (de forma, significat i funcionament sintàctic) en el corpus d'exclusió. Aquest criteri per detectar neologismes es diu **criteri lexicogràfic** i deixa la porta oberta a un nombre molt gran d'entrades que no interessin i, per això, a l'apartat 2.2 es presenten restriccions d'aquest criteri.

● A partir de l'any 2022, el **corpus d'exclusió** és **només** el diccionari normatiu en la versió en línia: DIEC (<<https://dlc.iec.cat/>>).

● Un dels canvis més importants de l'Observatori a partir de l'any 2022 és que fins ara un **neologisme** era una ocurrència d'un neologisme (tant si ja estava recollit a la nostra base de dades com si no ho estava) i ara un neologisme és el conjunt de totes les ocurrències d'una mateixa forma. A més, ara les variants s'agrupen (són neologismes diferents però estan internament lligats) i abans, no.

Pel que fa als **tipus de buidatge**, actualment, el buidatge pot ser manual, i les fitxes s'entren a la Plataforma Obneo, o automàtic, amb l'eina Buscaneo. I des del punt de vista dels textos buidats, el buidatge pot ser de premsa, de revistes, de textos radiofònics i audiovisuals, de xarxes socials i de textos especialitzats.

Hi ha diversos **grups** que fan buidatge en català: l'Observatori de Neologia i tots els grups de la xarxa d'observatoris de neologia de les varietats geolèctals del català (Neoxoc: <<http://www.iula.upf.edu/rec/neoxoc/>>).

Quan fem **buidatge manual**, marquem les paraules que podrien ser neològiques i després les busquem al diccionari. Els candidats que compleixen el criteri lexicogràfic són els que s'entren a la Plataforma, on passen per un procés de validació abans de ser publicats en el banc BOBNEO. I quan fem **buidatge automàtic** és el programa qui ens proposa una llista de candidats que nosaltres hem de triar i completar. A continuació expliquem el funcionament de la Plataforma i del Buscaneo.

2. Plataforma Obneo, per al buidatge manual

La Plataforma Obneo és l'eina que permet introduir les dades del buidatge manual, com acabem de dir, però permet també fer cerques i moltes altres tasques que van més enllà del buidatge. Quan s'hi accedeix (<<http://obneo.iula.upf.edu/pobneo/index.php>>) demana nom d'usuari i contrasenya perquè és una eina d'accés restringit als col·laboradors dels observatoris de neologia. A continuació s'expliquen els criteris per emplenar els camps de la fitxa que se'ns obriran quan entrem a la Plataforma i anem a l'apartat de gestió de fitxes i cliquem l'opció d'introduir noves fitxes.

El programa ens demanarà que indiquem la llengua, el tipus de buidatge, la font i la data del text de buidatge. A continuació, a la casella corresponent escriurem el neologisme que volem entrar i clicarem l'opció de buscar que hi ha al costat, de manera que el programa buscarà si aquella forma ja és al banc i ens oferirà els neologismes que trobi amb aquella paraula amb indicació de la categoria gramatical i el tipus de neologisme. Si la fitxa que volíem entrar es correspon amb algun d'aquests neologismes, el seleccionarem, el programa arrossegirà la informació de la categoria i el tipus, i entrarem la resta d'informacions en els camps previstos; en canvi, si no es correspon amb cap dels neologismes, o el programa no n'ha trobat cap, hauré d'emplenar tots els camps que se'ns obriran (que són lleugerament diferents segons el tipus de buidatge) i que s'expliquen a continuació.

2.1. Dades de la fitxa

Les dades d'aquest bloc es refereixen al neologisme i s'hi recullen aspectes d'anàlisi lingüística relacionats amb l'ús.

Entrada

En aquest camp entrem el neologisme en forma lematitzada tenint en compte els criteris següents:

1. No es consideren neològiques les paraules que ja apareixen en el DIEC (ja sigui amb entrada pròpia o en subentrada, en el cas de neologismes compostos o sintagmàtics). ● Si apareixen en un exemple, sí que els recollim i ho indiquem en una nota específica (nota 1).
2. No es consideren neologismes els noms propis, els diminutius, els augmentatius i els superlatius (si no són lexicalitzacions amb significat propi), els adverbis acabats en *-ment* i els mots prefixats amb *ex-*, ● *sub-*, *vice-* o *sots-* referits a càrrecs.
3. L'entrada s'enregistra en minúscules, excepte les sigles, si les trobem en majúscules.
4. Quan un mateix lexema es trobi sota dues categories gramaticals diferents (sobretot nom i adjectiu) es considera a tots els efectes que es tracta de dos neologismes diferents, i se'n fan dues fitxes, una per a cada neologisme. Per exemple, farem una fitxa per a *preadolescent* amb categoria *m i f* si trobem aquest neologisme en un context com "Aquesta música és apta per a *preadolescents*" i una altra per a *preadolescent* amb categoria *adj* si el trobem en un context com "Els joves *preadolescents* es dirigeixen cap a una parada de l'autobús".
5. Els neologismes repetits en una mateixa font de buidatge, només es recullen una vegada, però cal tornar-los a recollir quan aquests mateixos neologismes surten documentats en un altre buidatge (un altre dia o en una altra font).
6. ● Els substantius i els adjectius de doble terminació s'enregistren amb la forma lexemàtica, prescindint de la forma en què apareguin en el context. Així, en el context "La plataforma posarà en contacte els *desenvolupadors* d'aplicacions i el finançament per trobar solucions que resultin útils" recollirem la forma lematitzada *desenvolupador desenvolupadora* (i la categoria *m i f*); en canvi, en el context "En superar el curs, els candidats queden inscrits en una borsa de treball de les empreses *desenvolupadores*" recollirem la forma lematitzada *desenvolupador -a* (i la categoria *adj*).
7. Cal recordar que la forma femenina dels adjectius es consigna així: *contracíclic -a* (sense coma).
8. Els neologismes que només poden ser utilitzats en plural (plurals lexicalitzats) s'han d'entrar en forma de plural: *municipals* es consigna amb *f pl* en el context: "El Partit Popular presenta menys candidatures a les *municipals*".
9. Els verbs s'enregistren en infinitiu i en els verbs pronominals s'hi fa constar el pronom. Així, *enganxar-se* és un verb pronominal en el context: "Quan vas a casa d'algú que té tele, *t'hi enganxes*".
10. Els manlleus en què hi hagi dubtes en la lematització (per nombre o per gènere) es deixen en la forma en què apareguin al context.
11. Els adjectius que es corresponen amb el participi de passat d'un verb que no apareix en el diccionari s'han de considerar verbs i recollir-los en la forma d'infinitiu (si clarament només existeixen en la forma de participi, com *precuinat -ada*, sí que els recollirem com a adjectius).
12. ● Amb aquest nou protocol sí que recollim els gentilicis i les sigles de noms comuns (encara que apareguin escrites en majúscules).

Categoria gramatical

En aquest camp s'obre un menú desplegable on hem de seleccionar la categoria gramatical corresponent: *m, f, m i f, m pl, f pl, adj, adv, v tr, v intr, v pron, loc adj, loc adv, loc v, prep, conj, interj*.

Categoria		Comentaris	
m	nom masculí	És important observar en el context quin és el gènere de cada substantiu. Hi ha casos oscil·lants, que es poden documentar com a femení o com a masculí. En aquest cas farem una entrada per a cada forma de manera separada: el <i>covid</i> / la <i>covid</i> .	
f	nom femení		
m i f	nom masculí i femení	Per als noms d'oficis, càrrecs o membres de col·lectius (<i>anarcoecologista</i>). Recordeu desdoblar la entrada si les formes flexionades són diferents: <i>intendent intendenta</i> .	
m pl	nom masculí plural	Exclusivament per als neologismes que només poden documentar-se en masculí plural (<i>infrarojos, isquirotibials</i>).	
f pl	nom femení plural	Exclusivament per als neologismes que només poden documentar-se en femení plural (<i>agulletes, aigües obertes</i>).	
adj	adjectiu	Recordeu que si l'adjectiu presenta variants flexionades l'entrada s'hi representa de la manera següent: <i>abancalat -ada, abkhazi -àzia, abrahàmic -a</i> .	
adv	adverbi	Comproveu que la funció sigui clarament l'adverbial (i no l'adjectiva), tant si és un manlleu del castellà (<i>palante</i>) com de qualsevol altra llengua (<i>avanti</i>). Recordeu que no recollim els adverbis acabats en <i>-ment</i> .	
v tr	verb transitiu	Entrarem el verb neològic com a verb transitiu encara que en el context l'objecte directe estigui sobreenès, com en l'ús absolut (<i>anaves a mangar</i>) o funcioni com a subjecte gramatical en variants passives (<i>el document es pot baixar</i>), inacusatives (<i>no poden cabrejar-se</i>) o de participi (<i>els espais pacificats</i>).	
v intr	verb intransitiu	Recordeu que els verbs intransitius poden ser intransitius purs (<i>palmar, mailejar</i>) o de règim verbal (<i>sintonitzar amb, cristal·litzar en</i>). No confongueu el verbs intransitius amb els usos absoluts dels verbs transitius, com <i>mangar</i> o <i>pacificar</i> .	
v pron	verb pronominal	Reservat exclusivament per als verbs que <u>només</u> poden flexionar-se amb pronom: <i>autoinculpar-se, parapetar-se</i> . Les variants pronominals de verbs transitius o intransitius no s'han d'entrar com a <i>v pron</i> .	
conj	conjunció	Reservat per a casos simples no recollits al DIEC (<i>aviam, ergo, donques</i>) o per a combinacions del tipus <i>mentres que</i> .	
interj	interjecció	Reservat per a les onomatopeies o altres elements que funcionen en context com a interjeccions: <i>va dir uf, osti, què foto jo aquí</i> .	
prep	preposició	Reservat per a preposicions manlleuades (<i>chez, hasta</i>).	
loc adj	locució adjectiva	<ul style="list-style-type: none"> Reservat per als sintagmes preposicionals de funció adjectiva: <i>de baix cost, en 3D, de pro</i>. 	Hi ha algunes locucions que poden exercir tant la funció adjectiva com l'adverbial. Triarem l'opció que documentem en context. Cal parar atenció que quan un sintagma preposicional exerceix d'atribut d'un verb copulatiu o de complement predicatiu l'hem de considerar locució adjectiva: <i>ser de calaix, estar de morros, deixar de pasta de moniato</i> .
loc adv	locució adverbial	<ul style="list-style-type: none"> Reservat per als sintagmes preposicionals de funció adverbial: <i>a anys llum, de facto, de refiló, en antena</i>. 	
loc v	locució verbal	<ul style="list-style-type: none"> Reservat exclusivament per a frases fetes <u>de sentit metafòric en el context</u>, formades per un verb i un complement directe: <i>dir missa, llançar-se a la piscina, agafar aire, posar negre sobre blanc, obrir foc</i>. 	Quan un verb du un complement preposicional, encara que la combinació expressi un sentit metafòric, cal verificar si la metàfora la vehicula el sintagma preposicional, i si és així entrar com a neologisme el sintagma preposicional i considerar locució adverbial (<i>pagar o cobrar en negre</i>) o locució adjectiva (<i>deixar de la mà de déu</i>).

Tipus de neologisme

Camp tancat en què seleccionem el tipus de neologisme. És la part d'anàlisi que demana més atenció, especialment en l'entrada de neologismes nous. Per seleccionar el tipus de neologisme correctament cal conèixer la tipologia de neologismes que fa servir l'Observatori de Neologia que s'exposa a l'apartat 4 (Classificació de neologismes, per a tots els buidatges).

Neolosfera

● En aquest camp cal fer constar l'enllaç a la Neolosfera si és que aquell neologisme hi apareix.

Nota i codi nota

● Antigament hi havia moltes notes però ara només n'hi ha dues:

- **Nota 1:** indica que el neologisme apareix recollit com a **exemple** en el DIEC. Ex.: en el neologisme *bomba atòmica* posarem la nota: apareix al DIEC com a exemple en l'entrada "bomba".
- **Nota 2:** indica que el neologisme és una **variant ortogràfica** d'una paraula ja documentada en el DIEC: *marketing (màrqueting), tibi tibia (tebi tèbia)*. Ex.: variant ortogràfica de "màrqueting".

Diccionaris

● Com ja hem dit, el corpus d'exclusió està constituït únicament pel DIEC. En aquest camp es farà constar, a la fase de compleció, si els neologismes recollits apareixen documentats en les fonts lexicogràfiques següents:

Neoloteca	GDLC	GD62	
Cercaterm	DNV	VOX	DCVB

Context

El context és la frase que testimonia l'ús del neologisme en el text i es recull seguint aquests criteris:

1. Cada fitxa només admet un únic context. En cas que el neologisme aparegui més d'una vegada en la font de buidatge cal triar el context més adequat tenint en compte que es prefereixen els contextos definitoris o explicatius.
2. En el context s'ha de marcar el neologisme amb un asterisc al davant i un altre al darrere, sense cap espai en blanc. Si el neologisme apareix més d'un cop al mateix context, només es marca entre asteriscos la primera vegada.
3. Cal tenir en compte que si el neologisme apareix amb alguna marca tipogràfica com ara cometes o parèntesis, no cal reproduir-la en el context.
4. Idealment, el context ha d'incloure l'oració sencera en què apareix el neologisme: comença amb majúscules i es tanca amb un punt. Si l'oració és excessivament llarga i conté fragments que no aporten informació per a la comprensió del neologisme podem substituir-ne una part indicant-ho amb tres punts suspensius entre claudàtors ([...]), però cal assegurar-se que l'oració resultant té sentit ple.
5. Evitem recollir neologismes en contextos massa breus que no aporten informació sobre el neologisme. Però si el neologisme és molt interessant i no trobem un context més bo, podem recollir-lo encara que el context sigui insuficient.

Aspectes tipogràfics

Si en el context el neologisme apareix marcat amb cursiva, negreta o altres marques tipogràfiques caldrà indicar-ho en el desplegable (SÍ / NO).

Estat

● Una fitxa passa per quatre estats diferents i només apareixerà en el BOBNEO quan estigui validada.

Abans	Ara	
incompleta	pendent	mentre la persona que fa el buidatge encara hi treballa
completa	enviada	quan la persona que fa el buidatge ja l'ha considerat acabada
corregida	completa	quan la persona responsable del node ha omplert el camp dels diccionaris i blogs
validada	validada	quan la persona responsable del node valida i tanca la fitxa (en l'annex 2 hi ha la síntesi dels aspectes principals que cal tenir en compte en la validació)

Comentaris

Aquest camp només s'omplirà si la persona que fa el buidatge necessita consignar alguna informació rellevant no prevista a la fitxa.

2.2. Dades específiques del buidatge oral

Quan se selecciona el buidatge oral, cal emplenar els camps relatius a l'emissora, el programa i el tipus de programa.

En el camp de l'entrada, cal tenir en compte que no es recullen aquells castellanismes tan freqüents i propis del llenguatge oral espontani (*vale, pues, bueno, susto, aconteixement*, etc.) ni tampoc les variants orals del tipus *llavòrens, aixís, sapiguer*, etc. A més, en el buidatge oral cal omplir aquests camps específics:

- rol de l'emissor: per indicar quin tipus de paper desenvolupa el parlant que utilitza el neologisme:
 - conductor (responsable del programa, presentador o altre professional)
 - convidat (també col·laborador o tertulià)
 - públic
 - altres
- sexe de l'emissor: per indicar si el parlant és dona o home (hi ha també l'opció "no ho sé" per quan no es pugui determinar).
- dialecte de l'emissor: per indicar el dialecte del parlant (baleàric, central, septentrional, nord-occidental o valencià) i si no es pot determinar cal posar almenys si es tracta d'un dels grans dialectes (oriental o occidental).

2.3. Dades específiques del buidatge de Twitter i altres xarxes socials

A les xarxes socials com Twitter només recollim neologia lèxica, com en la resta de fonts; per tant, encara que sovint trobem formes que, malgrat no estar documentades als diccionaris, no corresponen a processos de creació lèxica, convé recordar que no les hem de recollir. Exemples: abreviacions com *q* (que), *aqt* (aquest) o *a10* (adéu); canvis de grafia com *kuan* (quan); repetició de lletres per indicar una determinada entonació (*moooooooooooooooooolt*), o combinacions de mots per formar etiquetes o *hashtags* (*#tuitserios*).

Igual com en el buidatge oral, tampoc no es recullen aquells castellanismes tan freqüents i propis del llenguatge espontani (*vale, pues, bueno, susto, aconteixement*, etc.).

Quant al context, en el buidatge de Twitter i altres xarxes socials seguim els mateixos criteris que en la resta de buidatges, de manera que mirem d'incloure l'oració sencera en què apareix el neologisme, començada amb majúscula i tancada amb un punt final (encara que al tuit original no hi siguin ni el punt final ni la majúscula inicial). I si el neologisme apareix sol o amb molt poc context i només s'entén perquè remet a una piulada anterior, recollim aquesta informació al camp *Comentaris* de la fitxa.

2.4. Dades específiques del buidatge especialitzat

Es desenvoluparà més endavant.

3. El Buscaneo, per al buidatge automàtic

El Buscaneo és l'eina de detecció automàtica de neologismes. Quan s'hi accedeix (<<http://obneo.iula.upf.edu/pobneo/index.php>>) demana nom d'usuari i contrasenya perquè és una eina d'accés restringit als col·laboradors de les xarxes d'observatoris de neologia.

El Buscaneo ens ofereix la llista de fonts (diaris o comptes de Twitter) començant pel més recent. Les que encara s'han de buidar tenen dos botons: un d'informació i un de candidats; en canvi, les que ja han estat buidades només tenen el botó d'informació (ja no hi ha candidats pendents). Cal seleccionar l'exemplar de la font que es vol buidar, clicar el botó de candidats i se'ns obrirà una pàgina que conté tots els candidats. Caldrà anar-los analitzant un per un, prement el botó "avaluar" (llevat que hi hagi candidats erronis, que es poden seleccionar i eliminar sense analitzar, perquè de vegades hi ha fonts que retornen molt de soroll). El programa arrossega les dades bàsiques (llengua, tipus de buidatge, font i data) i les enganxa a cada fitxa, però la selecció dels contextos i altres aspectes s'han de fer manualment.

El programa mostra tots els contextos del neologisme detectat. Si el candidat és bo, cal seleccionar el context més significatiu, clicar-lo i després prémer "acceptar candidat". Si el candidat no és bo (perquè el programa està agafant per error un fragment de text d'una altra llengua, o és un error de picatge o és un candidat rebutjable per altres motius (vegeu l'apartat 2.2), rebutjarem el candidat i cancel·larem l'operació si volem tornar enrere. Si acceptem el candidat, se'ns obrirà la pàgina següent, que ens permetrà completar els camps de la fitxa que el Buscaneo no pot emplenar automàticament.

Caldrà lematitzar bé el neologisme, seleccionar la categoria gramatical i emplenar la resta de camps de la mateixa manera que quan s'entra buidatge manual a la Plataforma (vegeu apartat 2.2).

4. La classificació de neologismes, per a tots els buidatges

Els neologismes es classifiquen segons els tipus següents:

derivació	PREF	prefixació
	SUF	sufixació
	PRSU	prefixació i sufixació
composició	COMP*	composició patrimonial
	CULT	composició culta
	HIBR	composició híbrida
	SINT	composició sintagmàtica
truncació	SIGL	siglació
	ACR	acronímia
	ABR*	abreviació
canvi	SEM*	semàntic
	SCG	subcategorial
	CONV	conversió
	LEX	lexicalització
	VAR	variació
manlleu	M*	manlleu
altres	A	altres processos

A continuació s'expliquen tots els processos. Cal tenir en compte que els asteriscos de la taula indiquen que el procés es subdivideix en diversos tipus.

La metodologia no permet classificar un neologisme en diferents processos i, de vegades, un mateix neologisme es podria classificar de diferents maneres. En aquests casos s'ha de prioritzar un procés seguint les indicacions que es destaquen en quadres com aquest.

4.1. Derivació

PREF Neologisme per prefixació

Neologisme format a partir de l'adjunció d'un prefix a una base (*reconsiderar, antitelevisió, hiperbò, proàrab*). Algunes preposicions i adverbis funcionen com a prefixos (*avant-, contra-, no-, etc.*) i, per tant, considerem PREF i no COMP les paraules que hi estan formades. ● La llista de prefixos més freqüents és aquesta (però convé consultar el DIEC perquè hi ha altres prefixos de caràcter especialitzat (com *sesqui-, ento-, eu-*) i amb freqüència baixa que també formen neologismes prefixats):

a- (an-)	contra-	extra-	meta-	poli-	sota-
ante-	des-	hemi-	micro-	post-	sots-
anti-	di-	hiper-	mini-	pre-	sub-
arxi-	dis-	hipo-	mono-	pro-	super-
auto-	en- (em-)	in- (im-, il-, i-)	multi-	pseudo-	supra-
avant-	endo-	infra-	neo-	quadri-	tetra- (tetr-)
bes-	entre-	inter-	no-	quasi-	trans- (tras-)
bi-	epi-	intra-	pan-	re-	tri-
circum-	es-	juxta-	para-	rere-	ultra-
cis-	ex-	macro-	peri-	semi-	uni-
con- (com-, col-, co-)	exo-	mega-	pluri-	sobre-	vice-

● Com ja hem dit en un apartat anterior, no cal recollir els neologismes formats amb els prefixos *ex-, sub-, vice-* o *sots-* quan s'adjunten a bases simples o derivades que fan referència a càrrecs.

SUF Neologisme per sufixació

Neologisme format a partir de l'adjunció d'un sufix a una base (*pànxing, rodoredià -ana, seleccionable*).

No recollim diminutius, augmentatius i superlatius, excepte si es tracta de lexicalitzacions [→ LEX], així com tampoc adverbis amb *-ment*, excepte si l'adjectiu no es troba al diccionari. ● En canvi, sí que es recullen els gentilicis, simples o compostos.

En el cas de neologismes nominals i adjectivals formats a partir del mateix sufix (*tanguista* m i f, adj; *dalinià -ana* m i f, adj) i no documentats a les obres lexicogràfiques, es consideren tots dos casos com a neologismes sufixats (i no l'un conversió de l'altre). Tanmateix, si el neologisme comporta un canvi de categoria (amb el mateix significat) respecte de l'entrada documentada a les obres lexicogràfiques, es considera que el procés de formació és la conversió sintàctica [→ CONV] i no pas la sufixació, com a *comunitari comunitària*_{m i f}, documentat com a adjectiu al DIEC.

PRSU Neologisme per prefixació o sufixació

El codi PRSU s'utilitza només si realment no es pot determinar l'ordre dels processos de prefixació i sufixació implicats: així, tot i que algunes paraules es poden analitzar de dues maneres possibles (per exemple, podem analitzar el substantiu *prejubilació* com [pre[[jubila]ció]], i aleshores seria PREF, o com [[pre[[jubila]ció]], i aleshores seria SUF), tenir en compte la semàntica resol la major part de dubtes. En aquests casos es recomana pensar com seria la definició del diccionari (● al marge que la base de la derivació sigui o no al DIEC):

PREF

autofinançar-se: 'finançar alguna cosa amb mitjans propis'
desestacionalitzar: 'deixar d'estacionalitzar'
indetenible: 'que no es pot detenir'
neothatcherisme: 'nou thatcherisme'
postconcili: 'després del concili'

SUF

autofinançable: 'que es pot autofinançar'
despigmentant: 'que despigmenta'
desaferrament: 'acció i efecte de desaferrar'
invisibilitzar: 'fer invisible'
postmodernitat: 'qualitat de postmodern'

Aquest criteri també és vàlid per classificar casos amb formants cultes [→ CULT, → HIBR]:

biosintètic -a	'relatiu a la biosíntesi'	→ SUF
nanotecnològic -a	'relatiu a la nanotecnologia'	→ SUF

Només en cas que no es pugui determinar quin és l'ordre de derivació, doncs, classificarem els casos com a PRSU i en la fase de validació se'n valorarà la pertinència.

4.2. Composició

● COMP Neologisme per composició patrimonial

Considerem COMP els neologismes formats a partir de dues bases (simples o complexes). Cal recordar que no hem de considerar bases els adverbis i les preposicions que funcionen com a prefixos [→ PREF].

Tenim quatre estructures de compostos i en cada neologisme haurem de seleccionar la que li correspon (i en els dos primers casos també la subestructura) que li correspon:

N+N ¹	coordinat (“que també és”)	<i>paquet bomba, teatre auditori</i>
	subordinat (“és un tipus” del primer o del segon N)	<i>escot banyera, mascle alfa</i>
A+A	vocal d'enllaç -i-	<i>blanquiblaui -ava, verd-i-blanc -a</i>
	vocal d'enllaç -o-	<i>publicoprivat -ada, maternoinfantil</i>
	sense vocal d'enllaç	<i>liberaldemòcrata, centrista-urbanita-liberal</i>
V+N		<i>aparcacòtxes, portapapers, atrapa-somnis</i>
N+A		<i>braguetallarga, galtavermell -a</i>
altres estructures		<i>capllaçar, totcamí, teva-meva</i>
escurçaments		<i>autocaravana, fotoperiodisme, bibliopiscina</i>
reduplicacions		<i>xunda-xunda, nini, parapapà</i>

Noteu que considerem COMP les paraules formades per escurçaments de paraules que funcionen autònomament a la llengua, com *auto* ‘automòbil’ a *autocaravana*, *foto* ‘fotografia’ a *fotoperiodisme* o *tele* ‘televisió’ a *teleaddicte* *teleaddicta*, de manera que queden diferenciats de neologismes com *fotodepilació* o *teletreballar*, que contenen els formants cultes *foto-* ‘llum’ i *tele-* ‘lluny’, respectivament i que, per tant, són neologismes formats per composició híbrida [→ HIBR] (consulteu l’annex 1 si teniu dubtes).

● CULT Neologisme per composició culta

Considerem CULT els compostos formats per una forma prefixada i una forma sufixada (*hortofrutícola*, *biocè*, *bibliòfag* *bibliòfaga*, *geopatia*).

Cal tenir present que alguns formants cultes coincideixen formalment amb mots patrimonials que provenen de truncacions i que formen COMP o HIBR (*aero-*, *agro-*, *auto-*, *biblio-*, *bio-*, *demo-*, *eco-*, *foto-*, *hetero-*, *homo-*, *narco-*, *petro-*, *radio-* i *tele-*). Podeu consultar l’annex 1 per veure en quins casos els neologismes amb aquests elements s’han de classificar d’una manera o d’una altra.

● HIBR Neologisme per composició híbrida

Considerem HIBR les paraules formades per un mot patrimonial i una forma prefixada (*euroescèptic* *euroescèptica*, *fotoprotector*, *cibercensura*) o per un mot patrimonial i una forma sufixada (*viatgeteca*, *serièfil* *serièfila*). La distinció entre prefix i forma prefixada es basa en la consideració que en fa el DIEC (a PREF hi ha la llista dels principals prefixos); quant a la diferència entre sufix i forma sufixada, cal consultar el GDLC, ja que recull tant els sufixos com les formes sufixades.

Consulteu l’annex 1 per veure en quins casos els neologismes amb formants cultes o que provenen d’escurçaments s’han de classificar com a CULT, HIBR o COMP.

SINT Neologisme per composició sintagmàtica

Neologisme format per una estructura sintàctica fixa: la majoria de neologismes sintagmàtics nominals tenen l’estructura de N+A o N+SP (*iogurt grec*, *piscina infinita*, *coca de vidre*), però també hi ha locucions sintagmàtiques (*picar pedra*, *a tota pastilla*).

¹ En alguns casos, darrere dels N+N hi ha compostos sintagmàtics (*paper clau*, *actriu estrella*), però en aquest protocol es classifiquen com a COMP.

4.3. Truncació

SIGL Neologisme per siglació

● Neologisme format per les lletres inicials d'un nom comú (que sol ser sintagmàtic) corresponents a una sigla, tant si s'escriuen en majúscula com en minúscula: *mena*, *pecé*, *api*, *uvi*, *oenagé*, *emapetrès*, *ADN*, *PCR*, *ERTO*. No recollim, però, les sigles corresponents a noms propis (*ONU*, *OMS*, *DIEC*, etc.).

Cal indicar en el desplegable que s'obre si es tracta d'una sigla lexicalitzada (*mena*, *oenagé*).

Tingueu en compte que els SIGL poden presentar característiques diferents, especialment pel que fa a la flexió de nombre i gènere: per exemple, podem trobar “els mena” o els “menes”.

Les sigles manllevades d'altres llengües s'han de classificar com a SIGL i no com a manlleus: *mp3*, *sms*, *dvd*, *nimby*.

ACR Neologisme per acronímia

Neologisme format per la combinació de segments de diverses paraules: *catanyol* (de *català* i *espanyol*), *sextorsió* (de *sexe* i *extorsió*), *horeca* (d'*hosteleria*, *restauració* i *cafeteria*).

● ABR Neologisme per abreviació

Neologisme format per abreviació. Cal seleccionar de quin tipus d'abreviació es tracta:

- **fonològica**: neologisme format per l'escurçament d'una paraula. Exemples: *prof_{mif}* (*professor* *professora*), *tele_f* (*televisió*), *micro_m* (*micròfon*), *prot_{mif}* (*protagonista*).
- **sintagmàtica**: neologisme format per l'escurçament d'un sintagma², tant si és del primer element (*motora_f* per *llanxa motora*, *municipals_{f pl}* per *eleccions municipals*, *farmacèutica_f* per *indústria farmacèutica*), com si és del segon element (*xarxa_f* per *xarxa social*, *caixer_m* per *caixer automàtic*). En tots dos casos es produeix també un canvi semàntic, però no ho indiquem amb cap nota.

Les abreviacions sintagmàtiques s'han de classificar, doncs, com a ABR i no com a CONV quan es nominalitza l'adjectiu d'un sintagma per abreviació (*privada_f* per *escola privada*) ni com a SEM quan l'element resultant de l'abreviació ja és al DIEC (*xarxa_f* per *xarxa social*).

4.4. Canvi

SEM Neologisme semàntic

Neologisme format per una modificació del significat d'una base lèxica (*sacsejar*, *edat d'or*, *epicentre*, *dinosaure*, *equació*, *carxofa*, *estat de gràcia*, *hipotecar*). La neologia semàntica sovint és difícil de detectar, perquè no presenta cap evidència formal d'aquesta modificació de sentit; cal identificar els neologismes sobretot per mitjà del context i comprovar-los amb cura en el diccionari.

També són neologismes semàntics els noms propis (de persona o marca registrada) usats com a noms comuns (*bollicao*, *vespa*, *xupa-xup*, *fitipaldi*).

● Cal indicar en el desplegable que s'obre si es tracta del canvi d'un nom propi a nom comú.

Playstation, *doodle* i altres noms propis que provenen sobretot de marques registrades angleses es consideren neologismes semàntics i no pas manlleus.

SCG Neologisme subcategorial

Neologisme que implica un canvi de subcategoria gramatical:

- Canvi de gènere d'una unitat nominal documentada al DIEC: *capellana* (vs. *capellà*), *arxipresta* o *arxiprestessa* (vs. *arxiprest*), *pitonís* (vs. *pitonissa*).
- Canvi de nombre d'una unitat nominal documentada al DIEC: *bermuda* (vs. *bermudes*).
- Canvi de règim verbal:

² De vegades pot ser un compost patrimonial (*voltor_m* per *fons voltor*).

- a) Variants transitives de verbs documentats al DIEC únicament com a intransitius: *telefonar* v tr (vs. *telefonar* v intr), *burlar* v tr (vs. *burlar-se* v intr pron), *contactar* v tr (vs. *contactar* v intr).
- b) Variants intransitives o de règim preposicional de verbs documentats al DIEC únicament com a transitius: *col·locar-se* v intr pron (vs. *col·locar* v tr), *cursar* v intr [en medicina, *una patologia cursa amb febre*, per exemple] (vs. *cursar* v tr).

Tingueu en compte que no considerem neologismes subcategorials com a canvi de règim verbal els usos absoluts de verbs transitius (*anaves a mangar*), ni tampoc les variants pronominals passives (*el document es pot baixar*), reflexives (*té poders per teletransportar-se d'un lloc a un altre*), recíproques (*aposta per coaliar-se amb l'oposició*), inacusatives (*no poden cabrejar-se*) o de participi (*els espais pacificats*) de verbs transitius.

LEX Neologisme per lexicalització

Neologisme format a partir de la lexicalització d'una forma flexiva, normalment del participi de passat del verb (*enquestat enquestada_{m i f}*). Així mateix, classifiquem com a LEX els derivats amb sufixos aspectius (diminutius i augmentatius) que tenen un significat no dimensional (*tovalloleta*, *gracieta*).

Quan el neologisme és un substantiu procedent d'un participi de passat lexicalitzat es considera sempre LEX, tant si aquesta forma verbal està documentada en el DIEC com si no ho està. En canvi, si funcionen com adjectius d'un verb documentat en el corpus d'exclusió es desestimen perquè es considera que queden recollides en el paradigma verbal corresponent (ex.: *il·lusionat -ada*, ja que *il·lusionar* està documentat). Si el verb no està documentat, es consideren neològics i es tracten d'acord amb el procés de formació que els ha generat (ex.: *redireccionat -ada* com a PREF sota l'entrada *redireccionar*).

Els adjectius morfològicament complexos que tenen com a base un participi verbal documentat com a adjectiu en el corpus d'exclusió, es tracten d'acord amb el procés que els ha generat (*precuinat -ada* i *sobreendeutat -ada* són PREF, ja que *cuinat -ada* i *endeutat -ada* ja estan documentats com a adjectius).

● En canvi, es consideren sufixats els neologismes amb *-nt* que fins a l'any 2021 consideràvem lexicalitzacions del participi de present (*il·lusionant_{m i f}*, *convocant_{m i f}*, *desequilibrant_{adj}*).

CONV Neologisme per conversió sintàctica

Neologisme format a partir d'un canvi de categoria gramatical sense modificació de la base lèxica (*transgènic_m* —documentat com a adjectiu—, *impagat_m* —documentat com a adjectiu—, *xipar_{v tr}* —es documenta *xip_m*—, *xof_{m i adj}* —documentat com a interjecció—).

VAR Neologisme per variació

Considerem exclusivament les variants menors (normalment de mots simples no construïts), normalment ortogràfiques i de vegades dialectals, d'una paraula documentada al DIEC: *gotim* (per *gotim*), *@dicte* (per *addicte*), *x-trem* (per *extrem*), *caler* (per *calé*), *abatzer* (per *esbarzer*), *fideuà* (per *fideuada*), *llangonissa* (per *llonganissa*), *sapiguer* (per *saber*), *xicolata* (per *xocolata*), etc.

En canvi, **no** considerem VAR:

- Les variants morfològiques (*acaminar* per *encaminar*, *grumós* per *grumollós*) que cal entrar a PREF o SUF [recordeu que la resta de variants s'han de classificar a SCG, CONV, S o el que pertoqui].
- Les oscil·lacions en l'ortografia dels compostos (*ràdio-oient* per *radiooient*, *llarg-metratge* per *llargmetratge*, *café-concert* per *café concert*), que cal ubicar a COMP, CULT o HIBR, segons correspongui, i sempre amb la nota 2 de variant ortogràfica.
- Les variants ortogràfiques i morfofonològiques de mots construïts (*esclar* per *és clar*, *sisplau* per *si us plau*), que recollim com a lexicalitzacions (LEX).
- Els simples errors ortogràfics o de picatge (*babor* per *bavor*, *gairabé* per *gairebé*, *garbuig* per *garbuix*), que no cal recollir com a neologismes, perquè són involuntaris.
- Els manlleus (*cannabis*, del llatí científic, per *cànnabis*, al DIEC; *expoli* per *espoli*), que cal classificar com a M.

- Les derivacions de mots formats per variació (*okupació d'okupa, xicolatada de xicolata*), que cal classificar com a SUF encara que la base no es reculli al DIEC.

Sempre, doncs, prioritzarem el procés de formació (SCG, PREF, SUF, COMP, etc., o M) i, quan escaigui, s'indicarà en la nota 2 que és una variant.

4.5. Manlleu

En aquest bloc recollim les paraules manllevades a una altra llengua (*heavy metal, after hour, rentrée, flygskam, sushi, gauche divine, amper, peixmerga*) i també els hel·lenismes i els llatinismes, que no són pròpiament manlleus (*modus operandi, dixit, pathos*).

☉ Distingim els manlleus segons la llengua de procedència, com s'indica en la taula següent. Si el manlleu no prové de cap de les llengües de la taula, li posem el codi MA (altres llengües). A més, tindrem en compte que:

- en el cas del grec, considerem hel·lenismes (HL) els cultismes clàssics i considerem manlleus els del grec modern (MGR);
- en el cas del llatí, no els considerem manlleus sinó llatinismes (LL).

hel·lenismes	HL	llatinismes	LL
--------------	-----------	-------------	-----------

alemany	MDE	castellà	MES	japonès	MJP
anglès	MEN	francès	MFR	portuguès	MPT
àrab	MAR	gallec	MGL	rus	MRU
asturià	MAST	grec	MGR	xinès	MCN
basc	MEU	italià	MIT	altres	MA

Marcarem el tic de manlleu adaptat només quan tinguem clar que la forma documentada no es correspon amb el manlleu original. Per fer-ho tindrem en compte el següent:

- per al castellà, prenem com a referència el DLE;
- per a la resta de llengües, podem consultar diccionaris de referència i, en cas de dubte, no marquem el neologisme com a manlleu adaptat;
- per a les llengües amb alfabet no llatí, només els considerem adaptats quan la forma no coincideix amb la forma transliterada (i, en cas de dubte, no marquem el neologisme com a manlleu adaptat).

Els manlleus poden ser simples, derivats, compostos, etc. Per norma general els considerarem manlleus, llevat de les sigles de manlleus (que etiquetem amb SIGL) i els noms propis manllevats usats com a noms comuns (que etiquetem amb SEM). Així doncs, classifiquem com a manlleus:

- els derivats construïts en la llengua original (*antiaging, macrobotellón*) [en canvi, *gamificar* i *sanchisme* són SUF perquè s'han construït en català sobre una base manllevada];
- els compostos construïts en la llengua original (*pop-rock, pàdel surf, bitxo raro, dolce vita*), encara que alguns dels elements estigui adaptat;
- els acrònims construïts en la llengua original (*docurreality, brèxit*) [en canvi, *docurealitat* és ACR];
- els escurçaments de manlleus sintagmàtics (*top, foie*).

4.6. Altres

Etiqueta reservada per a mots simples, que poden ser dialectals, argòtics, casos estranys, etc., que no apareixen als diccionaris (*carranco, marratxa*), i per a interjeccions i onomatopeies (*ups, ual·la*).

☉ També posarem A en aquells neologismes que no sapiguem com classificar per tal que es resolguin a la fase de validació.

En canvi, no registrarem com a A altres neologismes que podem classificar clarament en altres tipus ni tampoc no hi recollirem errors involuntaris.

Annex 1. Taula amb formants i truncacions que poden formar diversos tipus de compostos
 [tenint en compte el caràcter dinàmic dels formants, aquesta taula es revisarà periòdicament]

	CULT	HIBR	COMP
aero-	'aire'	'aire' aerodansa aeromarítim -a	'aeronau' aerolínia aeromecànica
agro-	'camp' agrobiologia agroecologia	'camp, agricultura' agroambiental agroramader -a	'=ecològic' agrobotiga agroturisme
auto-	'uno mismo' autocéfalo -la	*	'automòvil' autocaravana autoescuela
biblio-	'llibre' bibliòfag -a bibliometria	'llibre' biblioteràpia	'biblioteca' biblioesmorzar bibliopiscina
bio-	'vida' biocè	'vida' biociència bioètic -a	'biològic' biocosmètica bioplàstic
demo-	'poble' demoscòpia	'poble' demolingüística	'democràcia/democràtic' democristia -ana demosocialista
eco-	'medi' / 'so' ecògraf ecocardiograma	'medi' / 'so'	'ecològic' ecotaxa ecosostenible
euro-		'europeu' euroasiàtic -a eurocentrisme	'UE', '€' euroconversor eurocimera
foto-	'llum' fototip	'llum' fotodepilació fotoprotector	'fotografia' fotoperiodista fotorecepta
gastro-	'estómac' gastroenterocolitis	'estómac' gastrooperació	'gastronòmic' gastroturisme gastrocultural
hetero-	'altre'	'altre' heteroagressiu -iva heterodestructiu -iva	'heterosexual' heterofòbia heteronormatiu -iva
homo-	'mateix' homòleg homopatogen -ògena	'mateix' homoclima	'homosexual' homoeròtic -a homoparental
moto-		'motor' motoserra motonivelladora	'motocicleta' motoclub mototurisme
narco-	'son'	'son'	'droga' / 'narcotràfic' narcofossa narcopis
petro-	'pedra'	'pedra'	'petroli' petroislam petromonarquia
radio-	'radi' / 'raigs'	'radi' / 'raigs' radiofàrmac radioocultació	'radioactiu' / 'ràdio' radiofórmula radiocontrol
tele-	'distància' teledermatologia	'distància' teleassistència teletreballar	'televisió' teleaddicte teleconsum

* N. B. Casos como *autoconfinar-se* i *autoexigència* són PREF.

Annex 2. Aspectes que cal tenir en compte en la fase de validació de les fitxes

A la fase de validació, la persona responsable s'assegura que les fitxes s'han emplenat correctament i les marca com a validades. Un cop validades, es publiquen automàticament a BOBNEO i tots els usuaris les poden consultar.

És aconsellable que la validació es vagi duent a terme regularment, de manera que no s'acumuli un gran nombre de fitxes sense validar, i que es faci amb el nou protocol a mà i tenint obert el cercador de BOBNEO per poder consultar el conjunt dels dades si sorgeixen dubtes.

L'experiència ha demostrat que els errors es concentren en determinats camps que tractem a continuació. Així doncs, encara que la validació consisteix a revisar tots els camps, convé parar atenció especialment als punts següents:

1. Al camp de l'**entrada** cal comprovar si:
 - el neologisme està ben lematitzat i d'acord amb la categoria gramatical assignada i l'ús en el context (especialment cal fixar-se en la forma femenina dels adjectius i en el desplegament en dos gèneres dels noms que ho necessitin);
 - si és un neologisme pertinent (que no es tracti de neologismes que no s'han de recollir o, en el cas dels sintagmàtics, que no siguin sintagmes lliures).
2. Al camp de la **categoria gramatical** comprovem especialment que la categoria verbal sigui la correcta (transitiu, intransitiu o pronominal) i que els noms en plural siguin realment *pluralia tantum*.
3. En el **context** ens assegurem que el neologisme aparegui identificat entre dos asteriscos i que el context és correcte (llargària suficient, de punt a punt, i sense errors evidents).
4. En el camp de **tipus de neologisme** haurem d'assegurar que els neologismes estan ben classificats, especialment en els tipus següents:
 - els tipus que han patit canvis en el nou protocol: subdivisions a COMP, SEM i ACR, i redistribució de FCULT a CULT, HIBR i COMP;
 - els tipus on se solen acumular errors: LEX, CONV i VAR;
 - els tipus on s'acumulen els dubtes: A i PRSU.