

Getting informality into the dictionary: colloquial neologisms

ELISENDA BERNAL

Universitat Pompeu Fabra

MARTÍ FREIXAS

Universitat Pompeu Fabra

Abstract

For a language to be suited to all situations of use, its dictionaries must adequately represent colloquial words and expressions. Since the publication of the *Diccionari general de la llengua catalana* by Pompeu Fabra in 1932, lexicographical paradigms have changed significantly, as have perspectives on the various functional varieties each language possesses.

Thus, to prevent dictionaries from being seen as fossils only applicable to formal, but not colloquial, usage, colloquial lexis and phrasemes must be included. Two reasons apply: use of colloquial language is on the rise,¹ and the official language academy's new *Gramàtica de la llengua catalana* (2016) breaks new ground by including many colloquial structures and phrasings that do not appear in the normative dictionary, DIEC2. The lexis contained in this dictionary does not meet the needs of today's society, because we no longer limit our use of colloquialisms to informal situations. Rather, we tend to blend registers; the needs of dictionary users have diversified greatly.²

Including words and expressions from lower registers is all the more problematic considering the current state of most Catalan language dictionaries.³ Furthermore, current labelling systems that give usage information on pragmatics, register and speaker intentions are unsystematic and chaotic, demonstrating that colloquial lexis (and its inclusion in dictionaries) is one of the dimensions less attended by lexicographers.

¹ Margarida Bassols, 'El col·loquial dels mitjans', in Margarida Bassols, and Milà Segarra, eds, *El col·loquial dels mitjans de comunicació* (Vic: Eumo Editorial, 2009).

² Elisenda Bernal, Joan Costa, and David Paloma, *La codificació i l'estandardització del català avui* (Barcelona: Fundació per a la Universitat Oberta de Catalunya, 2010).

³ Jaume Salvanyà, 'Les marques de "baix" registre en els diccionaris', *Llengua Nacional* 72 (2010).

This study uses a corpus of colloquial lexis consisting of entries and subentries marked with an evaluative usage label in the four major Catalan language dictionaries, the DIEC2 (2007), the GDLC (1998), the DNV (2014) and VOX (2015), as well as in brief articles labelled *informal* on *ésAdir*, the language website of the state Catalan broadcasting corporation (CCMA). The 1,003-item corpus allows us, firstly, to describe and compare the way colloquial words contained in these dictionaries are treated through usage labels, with special emphasis on a comparison between the DIEC2 and the other works. Secondly, our lexicographical analysis is complemented with data generated by the NADIC project regarding neologisms labelled *col·loquial* [colloquial] in its database.

Our lexicographical analysis shows, among other things, that the boundaries between labels are poorly defined, and that lexicographers are themselves in disagreement about the limitations of usage labelling; contemporary lexicography will have to address this issue. Furthermore, we show that, while achieving global consensus on the formal/informal status of given words is relatively straightforward, the presence of connotations has the effect of creating various degrees of informality. Thus, we are led to complete a task begun rather timidly by Pompeu Fabra a century ago: defining a standard for colloquial language, as called for by *Secció Filològica*,⁴ in order to give colloquial Catalan the prestige it deserves in real everyday use.

Keywords

Catalan, colloquial language, usage labels, neology, dictionarization

⁴ *Secció Filològica*, ‘Sobre la variació i la prescripció en les orientacions normatives de la Secció Filològica’, in *Documents de la Secció Filològica*, V (Barcelona: Institut d’Estudis Catalans, 2018).

Introduction

The *Diccionari de la llengua catalana* (DIEC2) contains such uncommon words as *monopsoni* [monopsony, a situation in which a product only has a single buyer, who is able to set the market price], *pavesina* [small pavise], *sizígia* [syzygy, the point in the Moon's orbit in which it is in conjunction with or in opposition to the Sun] and *tómbol* [tombolo, a stretch of ground that connects an island with the coast or another island]. These words come from the fields of economics, warfare, astronomy and geography, respectively, and their use is extremely restricted, as attested in the *Corpus Textual Informatitzat de la Llengua Catalana* (CTILC):

Table 1. Occurrences and documents for specialist words in the CTILC

	occurrences/documents
<i>monopsoni</i>	—
<i>pavesina</i>	2/1
<i>sizígia</i>	4/2 (1 metalinguistic)
<i>tómbol</i>	2/2

However, the DIEC2 does not contain such commonly-heard words as *guai* [cool], *neura* [mania, obsession], *regatera* [‘cleavage’ or ‘coin slot’] and *xiripa* [luck], which are equally frequent, or more so, in the CTILC:

Table 2. Occurrences and documents for colloquial words in the CTILC

	occurrences/documents
<i>guai</i>	16/10 (1 metalinguistic)
<i>neura</i>	3/3
<i>regatera</i>	3/1
<i>xiripa</i>	19/13

The difference between these groups of words is that the former are specialist terms, which are well-represented in dictionaries, and the latter come from informal contexts, which have traditionally been less well covered. This is not a new situation. In 1951, Josep Pla, in *El carrer Estret* (chap. 33) lamented the lack of attention to colloquial

language: ‘La part que ocupa en la nostra vida de relació l’automatisme col·loquial és enorme, i és sorprenent com aquest no hagi promogut més interès.’⁵ [Conversational routine occupies an immense place in our social lives, and it is surprising that it has not awoken more interest. (Our translation)]

Years later, in 2016, philologist and translator Pau Vidal, in a seminar at the Universitat Pompeu Fabra, advocated the urgent creation of a “Col·locat”,⁶ to be modeled on TERMCAT (the official body regulating terminology in Catalan) but focusing on colloquial language. Dictionary users are disconcerted not to find words like *molar* [like, dig], *pixapins* [city dwellers who go to the countryside at the weekend] and *palla* [wank] in the normative dictionary: they do not understand the restrictions applied to some words. Indeed, attitudes toward colloquial language have changed immensely since the publication of *Diccionari general de la llengua catalana* (DGLC) in 1932.⁷ This work was tied to the values of the *Noucentisme* movement, and attempted to purify the language, after centuries of neglect, as well as shape citizens into ‘persones cultes, educades i de parlar polit’ [educated, polite and well-spoken people],⁸ even though it did not completely ignore colloquialisms:

És innegable que els paràmetres més ortodoxos del noucentisme imposaven una certa desconfiança respecte a la llengua popular i col·loquial —almenys en aquells girs i mots que es consideraven allunyats del bon gust—, però també cal afegir que aquí hi ha un problema tècnic que no és fàcil de resoldre encara avui: fins a quin punt un diccionari de la llengua estàndard ha de recollir mots que són aliens als àmbits d’ús formals.⁹

[It is undeniable that the most orthodox prescriptions of *Noucentisme* demanded a degree of mistrust of popular and colloquial language, at least when it came to turns of phrase and words considered outside the realm of good taste. In this situation there also lies a technical problem that is still difficult to solve: to what extent a dictionary of the standard variety of a language should include words that fall outside the bounds of formal situations. (Our translation)]

Despite the “technical problem” Ginebra and Solà mention, informality has come to receive significantly more attention, judging by references to *col·loquial* [colloquial], *familiar* [informal], *informal* [informal] and *popular* [popular] in the *Gramàtica de la*

⁵ Josep Pla, *El carrer Estret* (Barcelona: Destino, 1951).

⁶ The seminar was titled ‘*Es necessita Col·locat urgent*’ [‘Col·locat urgently needed’], and took place on 11 February 2016 as part of a monthly seminar series given by the InfoLex research group at the Universitat Pompeu Fabra.

⁷ Pompeu Fabra, *Diccionari general de la llengua catalana* (Barcelona: Edhasa, 1932).

⁸ Jordi Ginebra, and Joan Solà, *Pompeu Fabra: vida i obra* (Barcelona: Teide, 2007), § 5.56.

⁹ Ginebra, and Solà, *Pompeu Fabra: vida i obra*, § 5.57.

llengua catalana (GIEC, grammar of the Institut d'Estudis Catalans) as compared to the 1918/1933 and 1956 Fabra grammars:¹⁰

Table 3. Handling of colloquial registers in academic grammars

	Fabra 1918/1933 ¹¹	Fabra 1956	GIEC
<i>col·loquial</i> [colloquial]	2	2	124
<i>familiar</i> [informal]	1	—	2
<i>informal</i> [informal]	—	—	64
<i>popular</i> [popular]	—	—	2

Similarly, dictionaries have increased the number of colloquial words they contain, especially compared to the DGLC: currently, the DIEC2 contains 300 subentries labelled *popular* (although some of them appear to be obsolete or geographically restricted, like *andola* [peseta, the former Spanish currency], *betza* [paunch], *cassussa* [hunger] and *magriscolis* [thin]), forty-two labelled *vulgar* (all of them sexual in nature, like *quirra* [penis] and *boixar* [have sex]). However, the gap in coverage is still very large: the DIEC2 appears to maintain the overall approach set by the DGLC, in which Fabra aimed to establish the learned and formal registers while standardizing the language, whence the choice to create a prescriptive dictionary. As Bernal, Costa and Paloma¹² point out, while the DIEC2 has opened the door slightly to the inclusion of some words from lower registers, this is not enough to meet the needs of today's society, because these words are no longer restricted to informal situations, and user needs have become more diverse. For example, it is no longer enough to know whether *pipa* [sunflower seed] is in the dictionary or not, or if *guai* [cool] is allowed in some contexts; creators, translators, editors and other professionals feel the need for official guidance from the academic world on the use of these words.¹³

¹⁰ Jaume Salvanyà, 'El tractament de la llengua col·loquial: de Fabra als nostres dies', in *Andorra i Pompeu Fabra. Recull de les conferències sobre Pompeu Fabra* (Andorra la Vella: Centre de la Cultura Catalana, 2019), 39-60.

¹¹ In the 1918/1933 grammar, consulted on the *Portal Pompeu Fabra* (<www.upf.edu/web/portalspompeufabra>), the label *familiar* [informal] does not appear directly in the website's search engine, so data on it has been obtained from searching the PDF version of the work.

¹² Elisenda Bernal, Joan Costa, and David Paloma, *La codificació i l'estandardització del català avui* (Barcelona: Fundació per a la Universitat Oberta de Catalunya, 2010), 45-6.

¹³ Rudolf Ortega, 'El fill bord del Termet', *El País* (27 July 2015).

Towards such a proposal, we consider the data produced by the NADIC project (*Neologismes per a l'actualització del diccionari normatiu* [Neologisms for updating the normative dictionary]). However, it should be noted that the reason why the NADIC does not contain a large number of neologisms marked as colloquial is because they are provided by mass media with wide coverage (eminently written press),¹⁴ whose tone is rather formal, there is linguistic planning and creativity is restricted. Thus, all these factors define the natural usage of words that, in a way, deviate from the norm, e.g. colloquial, vulgar, slang, or even dialectal words.

For this reason, we wanted to complete the data from NADIC with a more detailed description of the treatment that colloquial words are given in various dictionaries, as well as the usage labels that usually accompany them. Both analyses (first one from NADIC and the other one from the study of normative dictionaries) should make it possible to demonstrate that colloquial lexis is not sufficiently represented in the main Catalan lexicographical works.

Lexicographical treatment of colloquial lexis

For this study, we formed a corpus of colloquial lexis, understood as entries (or subentries) which are marked as such by usage labels in the following dictionaries: the second edition (2007) of the *Diccionari de la llengua catalana* by the Institut d'Estudis Catalans (DIEC2), the *Gran diccionari de la llengua catalana*, 1998 edition, by Enciclopèdia Catalana (GDLC)¹⁵, the *Diccionari normatiu valencià* by the Acadèmia Valenciana de la Llengua, 2014 edition (DNV) and the third edition (2015) of the *Diccionari manual de la llengua catalana* by VOX (VOX).¹⁶

However, each of these dictionaries handles entries in different ways: in DIEC2 words and senses were identified with the usage labels *popular* [popular] and *vulgar* [vulgar];

¹⁴ Freixa, in this same book.

¹⁵ The GDLC (published in print format in 1998 and online at <www.diccionari.cat>) was regularly updated until 2010, and very sporadically until 2016. Since 2013, the permanently updatable version of the GDLC has been located at <www.enciclopedia.cat> and has been systematically updated since 2015. This is why some differences appear between one version and another. Our thanks to Núria Bort, editor of dictionaries at Enciclopèdia Catalana, for providing us with GDLC data.

¹⁶ Our thanks to Sofía Acebo, editor at Larousse, for providing us with data from the *Diccionari manual de la llengua catalana* (VOX).

in the GDLC, *col·loquial* [colloquial], *popular* [popular] and *vulgar* [vulgar]; in the DNV, *col·loquial* [colloquial] and *vulgar* [vulgar] and in VOX, *col·loquial* [colloquial] and *malsonant* [rude] (in some cases, usage is included in a footnote). To this list were also added the brief articles classified as *ús informal* [informal usage] on *ésAdir*, the language consultation website offered by CCMA (Corporació Catalana de Mitjans Audiovisuals), as well as the entries marked as colloquial in the NADIC project database (those included as well as those excluded).¹⁷ In total, the corpus contains 2,350 items, more than can be presented here, so the following filters have been applied to refine the list, if only in places, and reduce the count to 1,003:

- Only words from a major word class (noun, adjective, verb) were included. Further examples excluded are the formant *brom-*, labelled as vulgar in the GDLC,¹⁸ adverbs like *poder* [maybe] and *segurament* [probably] and the pronoun *en* (defined by the DIEC2 as ‘un adjectiu o substantiu indeterminat complement de ésser’ [an undetermined adjective or noun which is the complement of the copula *ésser* (our translation)]). Thus, expressions and set phrases were also excluded, e.g. *de pacotilla* [of poor quality], *estar tocat del bolet* [to be mad as a hatter], *gota ciàtica* [sciatica] and *partir peres* [lit. ‘split pears’: break up].
- Orthographic variants were also excluded: *antílop* (vs. *antílop*) [antelope], *music/musica* (vs. *músic/música*) [musician] o *tranyina* (vs. *teranyina*) [spider web].
- Items only appearing in the DNV were also excluded due to their limited geographical use: *embagassir* [to prostitute], *guisopo* [poor-quality stew], *birlonxo* [fat boy], *xeu* [from Alcoi].

Of the remaining 1,003 lexical units, on the one hand, all cases of polysemy (forty-two in total) were counted as separate entries, for example *flipar*, meaning both ‘go crazy for something’ and ‘get high’, and *súper*, meaning ‘supermarket’ as a noun and ‘great’ as

¹⁷ However, while collecting data for this study, we observed that the NADIC database contains other neologisms to which the same observation could have been applied. These examples include, especially, abbreviations (*apli* [app], *col·le* [school] *depre* [depressed], *físio* [physiotherapist], *mè/me* [shit], *paki* [Pakistani], *semis* [semi-final]) and castilianisms (*bronca* [scolding], *fardar* [show off], *guai* [cool] *marabunta* [infestation, crowd], *nòvio* [boyfriend], *nòvia* [girlfriend], *segurata* [security guard], *tonto -a* [silly], *xuleria* [cheekiness]).

¹⁸ According to the GDLC, the prefix *brom-* ‘bromo-’ is labelled as *obsolet* [obsolete] and *vulgar* [vulgar], although the latter label has nothing to do with the one we study in this work.

an adjective. On the other hand, a few examples of morphosyntactic variants were combined (*botellón/botellot* [drinking in the street], *òndia/òndima* [damn], *pixat/pixats* [pee]) as well as some units which were lemmatized differently between dictionaries (*atontar/atontat -ada* [stupefy], *folrar-se/folrat -ada* [get rich]).

All of these criteria are summarized in Table 4, which shows the number of entries with some sort of usage note or pragmatic information regarding register, and classifies them by work:

Table 4. Number of lexical entries analyzed in the corpus

		DIEC2	GDLC	DNV	VOX	ésAdir	NADIC
Usage notes	<i>col·loquial</i> [colloquial]		317	392	414		
	<i>malsonant</i> [rude]				36		
	<i>popular</i> [popular]	267	75				
	<i>vulgar</i> [vulgar]	42	108	50			
Other labels	<i>exclòs</i> [excluded]						2
	<i>inclòs</i> [included]						18
	<i>informal</i> [informal]					246	
		309	500	442	450	246	20

For our lexicographical analysis, entries occurring in more than one dictionary were combined, while maintaining the usage labels given in each dictionary of origin. The labels of *exclòs* [excluded] and *inclòs* [included] can only correspond with the NADIC corpus. Similarly, *informal* [informal] is the only label used by ésAdir for the lexical entries on their website, in accordance with the usage guidelines set by the site's editors (entries fall into categories like general use, unsuitable use, correct/incorrect use, etc.)

Below, we summarize the general characteristics of the dictionaries used.

Dictionaries published by the Institut

The DIEC2 labels colloquial words as *popular* [popular] and *vulgar* [vulgar], which is a reduction of the labels both in Fabra's DGLC (1932) and those in the DIEC1 (first

edition of Institut's *Diccionari de la llengua catalana*), with five and six categories, respectively, as shown in the following table:

Table 5. Usage labels in the Institut dictionaries

	DGLC	DIEC1	DIEC2	Examples
<i>familiar</i> [informal]	×			<i>endinyar</i> [to whack], <i>pendó</i> [libertine], <i>tiberi</i> [feast], <i>xaval</i> [kid]
<i>despectiu</i> [negatively connoted, pejorative]		×		<i>marieta</i> [lit. 'ladybird': poof], <i>menestralalla</i> [artisans], <i>verdulaire</i> [lit. 'greengrocer': rude]
<i>infantil</i> [child talk]		×		<i>bobò</i> [candy], <i>caca</i> [poo], <i>pipí</i> [pee], <i>pupa</i> [pain, injury], <i>titit</i> [bird]
<i>irònic</i> [ironic]	×	×		<i>matrona</i> [lit. 'midwife': fat woman], <i>patracols</i> [lit. 'thick books': paperwork], <i>predicaire</i> [lit. 'preacher': fraudster]
<i>poètic</i> [poetic]	×	×		<i>corser</i> [horse], <i>llir</i> [iris], <i>nuu</i> [naked], <i>ocella</i> [female bird]
<i>popular</i> [popular]	×	×	×	<i>afaitapobres</i> [lit. 'to shave + poor people': a nobody], <i>halar</i> [to eat], <i>nàpia</i> [nose]
<i>vulgar</i> [vulgar]	×	×	×	<i>collonada</i> [nonsense], <i>escalfabraguetes</i> [prick-tease], <i>magrejar</i> [to grope], <i>puteria</i> [slyness]

The DGLC also contains some words which Fabra did not mark with a label, but rather incorporated very general usage guidelines into the definition itself:¹⁹

advocadesc -a *adj.*

D'advocat, en mal sentit. *Habilitats advocadesques.*

[**lawyerly** *adj.*

¹⁹ Underlining ours.

Pertaining to a lawyer, in a negative sense. *Lawyerly abilities.*]

brau -ava *adj.*

Que excel·leix en el seu gènere (generalment en mal sentit). *És un brau lladre.*

[**good, skillful** *adj.*

Excelling in its category (generally in a negative sense). *She/he is a skillful thief.*]

càfila *f.*

Conjunt de gent, esp. en mal sentit, ardat. *Sou una càfila de mentiders.*

[**band** *feminine noun*

A group of gent, esp. in a negative sense, gang. *You are a band of liars.*]

The DIEC1 preserved this information and added the label *despectiu* [negatively connoted, pejorative]. The DIEC2 handled these units in an uneven fashion: for example, of the examples given just above, it only modified *advocadesc*, in order to harmonize it with other words with pejorative connotations (*invertit* [invert], *oficinesc* [bureaucratic], *xarnego* [a non-Catalan-speaking inhabitant of Catalonia with origins in another part of Spain]):²⁰

advocadesc -a *adj.*

D'advocat, dit despectivament. *Habilitats advocadesques.*

[**lawyerly** *adj.*

Pertaining to a lawyer, used pejoratively. *Lawyerly abilities.*]

However, it should be noted that changes to annotation strategies between dictionaries (expansion, reduction or combination of labels) does not necessarily mean a loss of information on usage, which is sometimes relegated to glosses, albeit less than systematically:²¹

arrencapins *m.*

Home menut i desnerit, dit irònicament.

[**shrimp** (lit. 'one who uproots pine trees') *masculine noun*

Short, stunted man, used ironically.]

bua *f.*

En el llenguatge infantil, mal. *T'has fet bua, pobreta?*

²⁰ Underlining ours.

²¹ Underlining ours.

[boo-boo *feminine noun*

In child talk, injury. *Do you have a boo-boo, poor thing?]*

cobrir-se de glòria *v. intr. pron.*

Irònicament, fracassar restant en ridícul.

[cover oneself in glory *intransitive pronominal verb*

Ironical. To fail and make a spectacle of oneself.]

histrió *m.*

Actor, especialment en sentit pejoratiu.

[ham *masculine noun*

Actor, especially in a pejorative sense.]

patuès *m.*

Parlar dialectal, especialment el privat de cultura literària i emprat sols en la conversa familiar, dit despectivament.

[patois *masculine noun*

Dialect, especially one lacking a literary tradition and used exclusively in familiar situations, used pejoratively.]

prou *adv.*

Usat irònicament, expressa negació. *Dius que és un bon home; sí, prou: ara t'explicaré quina bretolada m'ha fet.*

[quite *adverb*

Used ironically, it expresses negation. *You say he is a good man; yes, quite: now let me tell you about the dirty trick he played on me.*]

All these words appear with the corresponding usage label in the DIEC1, which contained a wider variety of labels (although this did not mean that all entries corresponding to an informal register or possessing expressive nuances were treated concisely and coherently). It is apparent that users need to be given guidance for using these words in a much more complete way than the two labels chosen for the DIEC2. By way of example, nineteen are labelled *infantil* [child talk] (*baba* [grandmother], *nyam-nyam* [yum yum], *titit* [little bird], etc.), eleven as *irònic* [ironic] (*bona peça* [little devil], *fer gràcia* [be funny], *saberut* [know-it-all], etc.), nine with as *despectiu* [negatively connotated, pejorative] (*advocadesc* [lawyerly], *descamisat* [lit. 'shirtless': shabby], *sequaç* [henchman], etc.) and only five as *pejoratiu* [pejorative] (*element*

[person], *histrió* [actor], *inqualificable* [indescribable], *lletrut -uda* [knowledgeable] and *de set soles* [lit. ‘of seven soles’: used as an intensifier]). It seems appropriate to take speaker intentions into account by adding information about these dimensions to the dictionary, as with examples of child talk. On the other hand, the difference between *pejoratiu* and *despectiu* does not seem to be clearly defined enough and should be revisited.

Comparison to other dictionaries

The way lexis is represented varies from dictionary to dictionary; this fact affects their microstructure. Each publisher applies a concrete set of standards for marking words with usage notes: in terms of pragmatics, this serves to signal register and speaker intentions when one form is chosen over another. As was discussed in § Lexicographical treatment of colloquial lexis, the other three dictionaries from which colloquial lexis was extracted for our corpus (the GDLC, the DNV and VOX) each offer a different treatment of usage labels from the DIEC2, as shown in Table 6:²²

Table 6. Usage labels encountered in the dictionaries studied

	DIEC2	GDLC	DNV	VOX
<i>col·loquial</i> [colloquial]		×	×	×
<i>malsonant</i> [rude]				×
<i>popular</i> [popular]	×	×		
<i>vulgar</i> [vulgar]	×	×	×	

This table shows that the DIEC2 is the only dictionary not to use the label *col·loquial* [colloquial]. On the other hand, it shares both *vulgar* [vulgar] and *popular* [popular] with the GDLC, which does include *col·loquial*. The data suggest that in the DIEC2,

²² Additionally, the *Gran diccionari 62 de la llengua* (2000) includes three usage labels different from those used in the other dictionaries studied here: *argot* [slang], *familiarment* [informal] and *pejorativament* [pejorative]. However, we were unable to include it in our analysis because it is only available in print format, making searches unwieldy.

popular corresponds to an extent with *col·loquial* in the GDLC (37.5 per cent of cases), as shown in Table 7:²³

Table 7. Correspondences between labelled entries in the DIEC2 and the GDLC

		GDLC												Total
		<i>col·loq.</i>		<i>pop.</i>		<i>vulg.</i>		D.L.		N.L.		N.E.		
		No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	
DIEC2	<i>pop.</i>	100	37.5	26	9.8	38	14.2	34	12.7	62	23.2	7	2.6	267
	<i>vulg.</i>	3	7.1	—	—	32	76.2	—	—	7	16.7	—	—	42

The table further shows that the use of *popular* overlaps very little between the two dictionaries (9.8 per cent), while *popular* (DIEC2) and *vulgar* (GDLC) do so slightly more (14.2 per cent). Another relevant fact shown in the table is that the GDLC often uses other labels for words marked as *popular* in the DIEC2 (12.7 per cent). Of these, the most frequently occurring is *despectiu* [negatively connotated, pejorative], in words like *homenot* [lit. ‘big man’: tomboy], *nyèbit* [rascal], *sastrinyol* [lit. ‘little tailor’: bad tailor]. Also frequent is *figurat* [figurative], in entries like *cagar-la* [lit. ‘shit it’: mess up], *carburar* [lit. ‘carburate’: work], and *trempera* [lit. ‘erection’: enthusiasm].

Lexicographically, the label *figurat* is not an evaluative term, but refers to a semantic transformation away from a literal sense, contained in unlabelled entries.²⁴ In some GDLC definitions labelled *figurat*, evaluative adjectives are included, intensifying the meaning. Some of these evaluative adjectives involve pejorative connotations (*desagradable* [unpleasant], *mal* [bad] *pitjor* [worse], etc.). This effect is also achieved with adverbs of manner ending in *-ment*, as in the following examples:²⁵

cromo *m*

fig. i col·loq. Persona excessivament acolorida, empolainada, vestida de coloraines, etc. *Vas fet un cromo.*

[lit. ‘**sticker**’ masculine noun

²³ The abbreviations in the right of the table refer to DIEC2 entries or subentries with a different label (D.L.), no label (N.L.) or no entry (N.E.) in the GDLC.

²⁴ Rosario González Pérez, ‘La marca *figurado* en los diccionarios de uso’, *Revista de Lexicografía* 7 (2000), 77–90.

José-Alvaro Porto Dapena, *Manual de técnica lexicográfica* (Madrid: Arco Libros, 2002).

²⁵ For a more detailed analysis of lexicographical usage labelling applied to derived words with evaluative suffixes, see Martí Freixas, *Els sufixos derivatius apreciatius en català: estudi morfopragmàtic i lexicogràfic* (Barcelona: Universitat Pompeu Fabra, 2018).

figurative and colloquial A person who is excessively colourful, dressed up, gaudy, etc. *You look like a sticker.*]

endrapar *v intr.*

fig. Menjar voraçment. *Endrapava com un lladre!*

[lit. ‘**cover with a cloth**’ *intransitive verb*

figurative Eat voraciously. *He was gorging himself like a thief!*]

mamarratxada *f.*

Acció extremament poca-solta, grotesca.

[**crackpot antic** *feminine noun*

An extremely scatterbrained or grotesque action.]

morrejar *v tr.*

fig. Besar intensament els llavis (d’algú).

[**snog** (li. ‘to snout’) *transitive verb fig.*

figurative Kiss someone’s lips intensely.]

genteta *f.*

desp. Gent roïna.

[**mob** *feminine noun*

pejorative Contemptible people.]

gramaticastre gramaticastra *m. i f.*

Gramàtic dolent, ignorant i pedant.

[‘**bad grammar**’ *masculine and feminine*

Bad, ignorant and pedantic grammarian]

politicastre politicastra *m. i f.*

desp. Mal polític.

[‘**slimy politician**’ *masculine and feminine*

pejorative Bad politician.]

Although the GDLC does not use a particular label to mark words from children’s speech, it does include an explanatory gloss within definitions, often with the phrase *terme infantívol* [term used by children] or *terme infantil* [term used by children]. Examples include the terms *meta* [breast] and *pipí* [pee]. As for terms marked *popular*

in the DIEC2, 23 per cent of them have no label or evaluative descriptor in their GDLC definition, including *pispar* [to steal] and *sonat -ada* [crackpot].

For the label *vulgar* [vulgar], used in both dictionaries, matches total 76 per cent of cases: of forty-two instances in DIEC2, thirty-two use the same label in the GDLC. This means that the same evaluative relationship holds in both dictionaries between the label and the entries it marks. It is essentially reserved for words related to sexuality, like *clau* [lit. ‘nail’: a screw], *colló* [ball] and *follar* [to shag]. Comparing the DIEC2 and the DNV yields similar results: for 267 of the words labelled *popular* [popular] in the DIEC2, 178 (66.7 per cent) are labelled *col·loquial* [colloquial] in the DNV:

Table 8. Correspondences between labelled entries in the DIEC2 and the DNV

		DNV										Total
		<i>col·loq.</i>		<i>vulg.</i>		D.L.		N.L.		N.E.		
		No.	%	No.	%	No.	%	No.	%	No.	%	
DIEC2	<i>pop.</i>	178	66.7	12	4.5	16	6	24	9	37	13.8	267
	<i>vulg.</i>	6	14.3	29	69	—	—	2	4.8	5	11.9	42

The difference lies in the dictionaries’ treatment of scatological terms (e.g., *merda* [shit] *merder* [place full of shit, mess], *tifa* [turd]), which, unlike sexual terms (labelled in both dictionaries as *vulgar*), are labelled *popular* in the DIEC2 and *vulgar* in the DNV, which explicitly indicates in its introduction that *vulgar* is to be used for words ‘considerats habitualment malsonants, grollers i poc o gens assumits pel bon gust majoritari’ [generally considered rude, offensive or unacceptable according to mainstream considerations of good taste (our translation)].²⁶

Lastly, VOX, like DIEC2, only uses two labels, *col·loquial* [colloquial] and *malsonant* [rude], so as to ‘no carregar en excés el text’ [avoid encumbering the text (our translation)], according to its user’s guide.²⁷ The *col·loquial* label, as with the other dictionaries, corresponds quite often to *popular* in the DIEC2 (47.6 per cent). On the other hand, the *malsonant* [rude] label, used only in VOX, is partially equivalent to *vulgar* [popular] in the DIEC2 (38.1 per cent):

²⁶ Acadèmia Valenciana de la Llengua, *Diccionari normatiu valencià* (València: Generalitat Valenciana, 2014).

²⁷ *Diccionari manual de llengua catalana* (3rd ed.) (Barcelona: Larousse, 2015), XVIII.

Table 9. Correspondences between labelled entries in DIEC2 and VOX

		VOX										Total
		<i>col-loq.</i>		<i>malson.</i>		D.L.		N.L.		N.E.		
		No.	%	No.	%	No.	%	No.	%	No.	%	
DIEC2	<i>pop.</i>	127	47.6	10	3.7	—	—	28	10.5	102	38.2	267
	<i>vulg.</i>	10	23.8	16	38.1	—	—	1	2.4	15	35.7	42

As with the DIEC2, VOX's two labels are sometimes insufficient to indicate contexts for use; thus, other complementary pragmatic information is frequently included as notes and observations, making use of other labels to nuance the treatment of a word's usage.²⁸

bemolls *m. pl.*

Collons. **NOTA** És un eufemisme de *collons*.

[lit. 'flats' (in music) *masculine plural*

Balls. **NOTE** Euphemism for *collons*.]

fatxa *m. i f. i adj.*

col-loq. Feixista. **NOTA** Té un valor despectiu.

[**facho** *masculine and feminine and adj.*

colloquial Fascist. **NOTE** Has a pejorative connotation.]

sidós -osa *adj.*

Que té sida. **OBS** Té un matís despectiu que fa preferibles altres formes com *afectat* o *malalt de sida*.

[**aids-ridden** *adjective*

Having aids. **OBS** Has a pejorative nuance, making the terms *afectat* [affected by] or *malalt de sida* [ill with aids] preferable.]

sisplau *m.*

Expressió de cortesia que es fa servir per demanar un favor. **OBS** Es pot fer servir en contextos que no siguin formals.

[**please** *masculine*

²⁸ However, compare VOX's entries for *gitanada* [gipsy trick] and *fornicar* [to fornicate], both labelled *despectiu* [negatively connoted, pejorative] and unaccompanied by any explanatory notes.

Politeness formula used to ask a favour. **OBS** Can be used in contexts that are not formal.]

teta f.

col·loq. Germana gran. **NOTA** És propi del llenguatge infantil.

[**big sis** *feminine*

colloquial Older sister. **NOTE** Comes from child talk.]

VOX's user's guide explains that the label *despectiu* [negatively connotated, pejorative] is applied to words that are not offensive but whose choice signals the speaker's contempt; still, it further specifies that in many cases 'és difícil que aquesta simple etiqueta reculli tot el que vol comunicar l'usuari'²⁹ [it is difficult for this mere usage label to encompass the full meaning which the user wishes to communicate (our translation)], which is why explanatory notes are sometimes added (especially in the case of insults).

To summarize, evaluative usage labels are not equivalent from one dictionary to the next, which is not in itself problematic. However, this situation leads to certain words receiving disparate lexicographical treatment, which produces some rather unexpected cases:³⁰

Table 10. Comparison of usage labels for different entries in Catalan dictionaries

	DIEC2	GDLC	DNV	VOX
<i>acollonir/acollonar</i> [intimidate]	<i>vulg.</i>	<i>vulg.</i>	<i>col·loq.</i>	<i>malson.</i>
<i>cagada</i> [shit]	<i>pop.</i>	<i>vulg.</i>	<i>col·loq.</i>	<i>col·loq.</i>
<i>camell</i> [dealer]	<i>pop.</i>	(<i>arg.</i>)	<i>col·loq.</i>	<i>col·loq.</i>
<i>clapar</i> [sleep]	<i>pop.</i>	<i>vulg.</i>	<i>col·loq.</i>	<i>col·loq.</i>
<i>cordons</i> [balls]	(<i>eufem.</i>)	<i>vulg./eufem.</i>	<i>col·loq.</i>	N.E.
<i>crac</i> [crack cocaine]	N.L.	(<i>arg.</i>)	N.L.	<i>col·loq.</i>
<i>emprenyar</i> [bother]'	<i>pop.</i>	<i>vulg.</i>	<i>col·loq.</i>	<i>malson.</i>
<i>figaflo</i> [foolish]	<i>pop.</i>	<i>desp.</i>	<i>col·loq.</i>	N.L.

²⁹ *Diccionari manual de llengua catalana*, XIX.

³⁰ Some of the dictionaries consulted include usage information within the definition as an explanatory gloss; the table includes this information between rounded brackets, alongside the corresponding pragmatic value. However, if the entry appears without a label, 'N.L.' is used (or 'N.E.' if there is no entry).

<i>gavatx -a</i> [French person]	N.L.	<i>pop./desp.</i>	<i>pej.</i>	<i>col·loq./ (desp.)</i>
<i>merda</i> [shit]	<i>pop.</i>	<i>vulg.</i>	<i>vulg.</i>	<i>col·loq.</i>
<i>milhomes</i> [conceited]	<i>s/m</i>	<i>iròn.</i>	<i>iròn.</i>	<i>col·loq.</i>
<i>putada</i> [dirty trick]	<i>pop.</i>	<i>vulg.</i>	<i>col·loq.</i>	<i>malson.</i>
<i>sastrinyol</i> [bad tailor]	<i>pop.</i>	<i>desp.</i>	<i>pej.</i>	N.E.
<i>titola</i> [penis]	<i>pop.</i>	<i>col·loq.</i>	<i>vulg.</i>	<i>col·loq.</i>
<i>trempera</i> [erection]	<i>vulg.</i>	<i>col·loq.</i>	<i>vulg.</i>	<i>col·loq.</i>

Given the data above, we can draw the following conclusions regarding the meanings associated with the usage labels used in these four dictionaries: in the DIEC2, *popular* [popular] correlates fairly often to *col·loquial* in the other dictionaries. In fact, half of the words marked *popular* in the former (50.6 per cent) are marked *col·loquial* in other works. Entries labelled *vulgar* in the DIEC2 have the same label in other dictionaries as much as 44.4 per cent of the time (given that VOX uses the equivalent label *malsonant* [rude] instead).

Variance in choices regarding labels, and especially in the way they are assigned, suggests the influence of ideology on dictionary design. While, in revising entries, the DIEC2³¹ made every attempt to attain maximum objectivity and neutrality in entries related to religion or questions of gender, informal lexis still needs in-depth study to improve the way it is currently handled. In general, apart from the words it contains that are no longer current, the DIEC2 labels sexual terms as *vulgar*, while scatological ones are labelled *popular*. However, while more or less unconscious moral value judgments may be at work, it is also true that the boundaries between different labels are poorly defined, and even lexicographers disagree about the limits of labels marking pragmatic information,³² which constitutes a challenge that today's lexicography will have to face. Garriga³³ already pointed out in the 1990s that lexicographical practice relies on intuition when it comes to assessing usage; despite advances in corpus techniques, this aspect depends in large part on the lexicographer's linguistic competence. As a result, usage labels are one of the '[aspectos] más faltos de sistematización en la lexicografía

³¹ Institut d'Estudis Catalans, *Diccionari de la llengua catalana* (2nd ed.) (Barcelona: Edicions 62; Enciclopèdia Catalana, 2007), introduction, § II.4.

³² Jaume Salvanyà, 'Les marques de "baix" registre en els diccionaris', *Llengua Nacional* 72 (2010), 16–19.

Kory Stamper, *Word by Word: The Secret Life of Dictionaries* (New York: Pantheon Books, 2017).

³³ Cecilio Garriga, 'La marca de "vulgar" en el DRAE: de Autoridades a 1992', *Sintagma* 6 (1994), 5.

actual'³⁴ [most lacking [aspects] in current lexicography (our translation).]³⁵ We will return to this question below.

Contribution of the NADIC project

The NADIC results, which started with 2,462 neologisms, crystallized as a proposal to include 927 words (distributed between new lemmas and new senses or subentries).³⁶ Of these, eighteen are labelled as colloquial in the database (1.94 per cent of the total new words).³⁷ These words can be classified into two major groupings: abbreviations and borrowings of Spanish or English words (mostly Spanish):

- *bici* [bike], *crono* [chronometer], *disco* [discotheque], *hetero/hétero/hètero* [heterosexual], *mani* [demonstration], *micro* [microphone], *narco* [drug dealer], *pel·li* [film], *progre* [progressive], *químiu* [chemotherapy], *retro* [retro], *súper* [supermarket] and *tele* [TV];
- *birra* [beer], *descollonar-se* [to crack up], *flipar* [to go crazy for; to get high], *freaky/friqui /friki* [freaky] and *guiri* [foreign tourist].

All the words in this list meet a minimum of three of the criteria established by the project to identify good candidates for dictionary additions.³⁸

Table 11. Distribution of criteria for including neologisms labelled colloquial [colloquial]

Usage				Documentary					Ling.	Total criteria
U1	U2	U3	U4	D1	D2	D3	D4	D5	L1	

³⁴ Garriga, 'La marca de "vulgar" en el DRAE: de Autoridades a 1992', 5.

³⁵ Similarly, Payrató criticizes the fact that the label *popular*, which is difficult to understand, has established itself so strongly in dictionaries, despite having little justification and being so close to the labels *vulgarism* [vulgarism] and (*propi del llenguatge*) *infantil* [(typical of) child language] (Lluís Payrató, 'Lexicografia i variació funcional', *Caplletra* 17 (1994), 47–58.)

³⁶ Freixa, in this same book.

³⁷ Two more words were excluded from the proposal because they did not meet set criteria (*escaquejar-se* [to slope off] only meets the criterion of being attested in complementary dictionaries; *putejar* [to screw over] meets two, attestation in both priority and complementary dictionaries).

³⁸ The usage criteria (Freixa, in this same book) were high frequency (U1), very high frequency (U2), support by project specialists (U3) and orientation towards usage (U4). Documentary criteria were existence of a TERMCAT proposal for normalization (D1), appearance in priority dictionaries (D2), dictionaries of other languages (D3), complementary dictionaries (D4), and in corpora. (D5). The linguistic criterion, in our corpus, was unpredictability (L1).

<i>bici</i>	×				×	×	×				4
<i>birra</i>	×		×					×			3
<i>crono</i>	×	×					×		×		4
<i>descollonar-se</i>	×		×					×			3
<i>disco</i>						×	×	×			3
<i>flipar</i>	×					×		×			3
<i>freaky</i>	×	×						×			3
<i>guiiri</i>	×					×		×			3
<i>hetero</i>	×						×			×	3
<i>mani</i>	×							×	×		3
<i>micro</i>						×	×	×			3
<i>narco</i>	×	×				×	×	×			5
<i>pel·li</i>	×	×		×							3
<i>progre</i>	×		×			×					3
<i>químio</i>	×		×				×				3
<i>retro</i>	×				×	×		×			4
<i>súper</i>	×					×		×			3
<i>tele</i>	×	×				×	×	×			5

Table 11 shows how the majority of entries share several criteria, especially those of high frequency (sixteen neologisms), presence in complementary dictionaries (twelve neologisms) and priority dictionaries (ten neologisms).³⁹ Attestation of a colloquial word in at least two dictionaries of other languages (*Diccionario de la lengua española*, *Lo Zingarelli* and *Le Petit Robert*) constitutes a criterion because it demonstrates that the word has already been sufficiently vetted to be included in a normative dictionary. (Indeed, *bici* [bike], *disco* [discotheque], *micro* [microphone], *narco* [drug dealer] *retro* [retro] and *tele* [TV] meet three of the five documentary criteria.)⁴⁰ What is clear, however, is that these words only represent a small, well-defined part of colloquial lexis

³⁹ The priority dictionaries (D2) are the GDLC, the *Gran diccionari 62 de la llengua catalana*, VOX and the *Diccionari d'ús dels verbs catalans*. The complementary dictionaries (D4) were the *Diccionari català-valencià-balear*, the DNV, ésAdir, the blog *Silencis del DIEC* and the *Gran enciclopèdia catalana*.

⁴⁰ As for formation processes, the NADIC data largely match the lexis with usage guidelines contained on the website ésAdir. Many of these words, which have also served to expand the corpus, are truncations (*boli* [pen], *furgo* [pickup truck], *maria* [marijuana], *mates* [maths], *secre* [secretary]) and borrowings from Spanish (*bolso* [handbag], *cutre* [stingy, shabby], *mandanga* [calm], *quinqui* [delinquent] and *rasca* [cold]).

in the Catalan language: abbreviations.⁴¹ Colloquial lexis included in a dictionary should not be limited to one type of formations;⁴² the spoken registers are far more rich and diverse than that.⁴³

Conclusion (or desiderata)

In view of the lexicographical analysis about the treatment of informality, we can confirm some of the objectives that had been outlined: colloquial lexis has been much less attended by lexicographers and this has meant that the presence in dictionaries is very low. Consequently, this lack of attention has affected the representation of this type of lexis in these works. In addition, the existing problem around the usage labels in the same dictionary and among the other dictionaries shows once again the lack of a clear, homogeneous and systematic proposal.

As for the dictionarisation of these new words in Catalan, the NADIC data are a first approximation to this kind of colloquial words that meet the most documentary criteria to be included in a normative dictionary: abbreviations and borrowings of Spanish or English words. Despite this, according to Salvanyà,⁴⁴ colloquial lexis (and phrases) must necessarily be included in dictionaries for two reasons: firstly, to increase the coverage of colloquial language;⁴⁵ secondly, because the Institut's official grammar, the GIEC, has made progress in this direction by incorporating numerous colloquial structures and phrases, which are nonetheless absent from the dictionary. Examples include *supermotivats -ada* [highly motivated],⁴⁶ *arxiconegut* [a household name],⁴⁷

⁴¹ Institut d'Estudis Catalans, *Gramàtica de la llengua catalana* (Barcelona: Institut d'Estudis Catalans, 2016), § 6.5.3.

⁴² With a project like NADIC, the vast majority of sources are semi-formal mass media, which explains the low incidence of other kinds of formations.

⁴³ Elisenda Bernal, and Carsten Sinner, 'Al seu rotllo: aproximació al llenguatge juvenil català', *Zeitschrift für Katalanistik* 22 (2009), 7–36.

Anna Matamala, 'Del guió a la pantalla: trets de l'oral col·loquial en la ficció televisiva', *Caplletra* 44 (2008), 141–168.

⁴⁴ Salvanyà, 'El tractament de la llengua col·loquial: de Fabra als nostres dies'.

⁴⁵ Margarida Bassols, 'El col·loquial dels mitjans', in Margarida Bassols, and Milà Segarra, eds, *El col·loquial dels mitjans de comunicació* (Vic: Eumo Editorial, 2009), 11–20.

Òscar Bladas, 'La llengua oral col·loquial', in Margarida Bassols, and Milà Segarra, eds, *El col·loquial dels mitjans de comunicació* (Vic: Eumo Editorial, 2009), 23–39.

⁴⁶ Institut d'Estudis Catalans, *Gramàtica de la llengua catalana*, § 11.6.a.

⁴⁷ Institut d'Estudis Catalans, *Gramàtica de la llengua catalana*, § 11.6.b.

rumorologia [gossip]⁴⁸ and even *xungo -a* [crappy].⁴⁹ The GIEC has truly made a qualitative leap by explicitly including functional variation.⁵⁰

[El model propi de l'Institut d'Estudis Catalans] [É]s un model que admet la variació pel fet que concep la llengua com un conjunt d'usos propis i adequats segons les circumstàncies geogràfiques, socials o funcionals en què es donen aquests usos. La variació geolectal existeix en totes les llengües, sigui quina sigui llur extensió geogràfica. Tots els parlants, sense excepció, s'han de poder sentir representats en els fenòmens recollits a la gramàtica, i també en les unitats que conté el diccionari. [...]

És un model que preveu la variació funcional o de registre, és a dir, que parteix de l'evidència que els parlants seleccionen usos diferents d'acord amb les característiques de cada situació comunicativa, ja sigui per raó del canal (escrit/oral) o ja sigui per raó del grau de formalitat que una situació requereix (formal/informal).⁵¹

[[The Institut d'Estudis Catalans' model] is a model that allows for variation due to the fact that it conceptualizes language as a set of particular uses that are adequate in terms of the geographical, social and functional circumstances in which they occur. Geolectal variation exists in all languages, whatever the size of their territory. All speakers, without exception, need to be able to feel represented in the phenomena included in the grammar, as well as the lexical units contained in the dictionary. [...]

This model anticipates functional or register variation, i.e., it starts with the obvious point that speakers choose different uses in keeping with the characteristics of each communicative situation, whether they vary by channel (written/spoken) or by the degree of formality required by a situation (formal/informal). (Our translation)]

Furthermore, Secció Filològica, in setting out the criteria followed in normative guidelines, has the following to say:

Tot això fa pensar en la necessitat de distingir, al marge dels registres formals de l'estàndard, unes característiques específiques d'un estil estàndard col·loquial —que alguns anomenen col·loquial mediatitzat o col·loquial de ficció—, a fi de concretar quins trets poden considerar-se col·loquialismes generals —gerundis en *-guent* (*poguent*), l'article neutre *lo* (*lo més estrany és això*) o determinades construccions sintàctiques (*Ets el responsable de que hagi caigut*, *La veïna que el seu marit és mecànic*)— i quins s'han de considerar formes col·loquials exclusivament dialectals —*sapiguer* (o *sebre*), *dugues* (numeral), *vem* (per *vam*), *estodiar*, etc.— o bé interferències d'altres llengües.⁵²

[All of this points to the need to distinguish, beyond the bounds of formal and standard registers, specific characteristics of a standard colloquial style (called 'media colloquial' or 'fiction colloquial' by some) in order to determine which features could be considered general-use colloquialisms (gerunds in *-guent* like *poguent* [being able], the neuter article *lo*: *lo més estrany és això* [the strangest thing is this] and certain syntactic constructions: *Ets el responsable de que hagi caigut* [You are responsible for it falling], *La veïna que el seu marit és mecànic* [The neighbour whose husband is a mechanic])⁵³ on one hand, and on the other, which colloquial forms should be considered exclusively dialectal (*sapiguer*, *sebre* [to know], *dugues* [two], *vem* (for *vam*, the auxiliary used in the periphrastic past

⁴⁸ Institut d'Estudis Catalans, *Gramàtica de la llengua catalana*, § 12.7.

⁴⁹ Secció Filològica, 'Sobre la variació i la prescripció en les orientacions normatives de la Secció Filològica', in *Documents de la Secció Filològica*, V (Barcelona: Institut d'Estudis Catalans, 2018), 45.

⁵⁰ Underlining ours.

⁵¹ M. Teresa Cabré, 'De la descripció a la norma: el model de llengua de l'Institut d'Estudis Catalans', in *Gramàtica de la llengua catalana* (Barcelona: Institut d'Estudis Catalans, 2016), XIV.

⁵² Secció Filològica, 'Sobre la variació i la prescripció en les orientacions normatives de la Secció Filològica', 41.

⁵³ In standard written Catalan, these examples would be '*Ets el responsable que hagi caigut*' and '*La veïna el marit de la qual és mecànic*'.

tense) *estodiar* [study], etc.) as well as interferences from other languages. (Our translation)]

Although the document appears to be addressing mostly morphosyntactic aspects, it later⁵⁴ recognizes lexical variation in informal registers:⁵⁵

[...] també dins dels registres informals hi pot haver trets de llenguatge infantil —*bobò*—, argòtic —*xungo*, *endrapar*— o vulgar —*emmerdar*.

[[...] also in informal registers there can appear features of child speech (*bobò* [boo-boo]), slang (*xungo* [crappy], *endrapar* [gorge oneself]) and vulgar terms (*emmerdar* [screw up]). (Our translation)]

Thus, lexis needs to be included in this variation in order to prevent the dictionary from being perceived as an artifact only relevant to solemn pronouncements but not to colloquial speech. To do so, we need to open the door to the use of primary documents as sources, as the *Oxford English Dictionary* does.⁵⁶ Furthermore, as mentioned above, we need to face the task of developing a labelling system that is more transparent and, as far as possible, more objective, in order to overcome the lacks in current systems. For example, *despectiu* [negatively connotated, pejorative] seems to be related to the concept of contempt, while *pejoratiu* [pejorative] is associated with an unfavorable evaluation. But probably all speakers have the perception that contempt is itself an unfavorable evaluation... The task of defining and establishing the boundaries of these labels is, without a doubt, difficult, but doing away with them is not a satisfactory answer, either, because a way of signaling nuances would be lost. Some proposals have been made in this direction, like Salvanyà,⁵⁷ who advocates a single umbrella term, *col·loquial* [colloquial], to cover all these uses. For Salvanyà, this change would present at least three advantages:

- a) It is a rather precise term, serving to restrict the use of the words it is applied to, but without emitting value judgements which are, at least in part, subjective.
- b) It is simple and easily implemented.

⁵⁴ Secció Filològica, ‘Sobre la variació i la prescripció en les orientacions normatives de la Secció Filològica’, 45.

⁵⁵ Underlining ours.

⁵⁶ In December 2018, the editors of the *Oxford English Dictionary* published an infographic on the inclusion criteria used for adding entries (<<https://public.oed.com/how-words-enter-the-OED/graphic/>>). A point that stands out is the diversity of sources used for confirming the use of a word: beyond the traditional and inevitable textual corpora and the media, the editors also consult online forums, television series, advertising, academic studies, Twitter and teen magazines.

⁵⁷ Salvanyà, ‘Les marques de “baix” registre en els diccionaris’.

- c) It would bring needed reform and enable the abandonment of mutually overlapping labels, which end up disorienting dictionary searches rather than guiding them.

Is it enough to use only one usage label? Is a binary opposition between *formal* and *informal* (actually ternary, given the middle term *neutre* [neutral]) sufficient for guiding speakers in the use of certain words? We consider the distinction between formality and informality to be an easy object of consensus; however, it is also true that there do exist connotations that influence choices: there are degrees of informality. For example, speakers intuit differences between words like *fantasmada* [boasting] (rather neutral), *cagarrines* [the runs] (somewhat childish) *pispar* [to steal] (colloquial) and *polla* [cock] (vulgar). Nevertheless, it is necessary to overcome our squeamishness surrounding ‘temes considerats tabú com el sexe i les funcions fisiològiques’ [topics considered taboo, like sex or bodily functions (our translation)], as was already advocated in López del Castillo.⁵⁸ The normative dictionary must cease to be seen as a receptacle for high registers by including more diverse and ‘current’ uses⁵⁹: in many cases, these uses are stabilized and unlikely to recede.⁶⁰

In a nutshell, the idea is to complete a process timidly started by Fabra (and which other dictionaries have continued) with words like *calé* [money], *esmorrar-se* [to fall down] and *xamba* [luck] (respectively, a borrowing from the Calo language, a word formed by prefixing and a borrowing from Spanish). Dictionaries have admitted words like *molar* [like], *esconyar* [smash up] and *xulo -a* [nice], analogous to the preceding ones, and agreed, as Stamper⁶¹ wrote about the *Merriam-Webster Dictionary*, that dictionaries must contain not merely ‘good’ words, but also ‘bad’ ones and ‘ugly’ ones (if these adjectives can truly be applied to words: beauty is in the eye of the beholder!) provided

⁵⁸ Lluís López del Castillo, *Llengua standard i nivells de llenguatge* (Barcelona: Laia, 1976), 49.

⁵⁹ Remember García Márquez’s 1981 criticism of the *Diccionario de la Real Academia Española* [Dictionary of the Royal Spanish Academy]: ‘En el diccionario de la Real Academia de la Lengua [*sic*], en cambio, las palabras son admitidas cuando ya están a punto de morir, gastadas por el uso, y sus definiciones rígidas parecen colgadas de un clavo.’ [On the other hand, words are admitted to the Real Academia de la Lengua [*sic*] dictionary when they are on the point of dying, worn from use, and its rigid definitions seem to be hung from a nail. (Our translation)] (Gabriel García Márquez, ‘La mujer que escribió un diccionario’, *El País* (10 February 1981).)

⁶⁰ Solà warned about the risks involved when the linguistic norm attempts to ignore certain structures that are well-established, even if they have been considered spurious for many years: ‘la normativa haurà de ser realista, si no vol ser olímpicament ignorada’ [the norm must be realistic, if it does not want to be outright ignored (our translation)]. (Joan Solà, *Sintaxi normativa: estat de la qüestió* (Barcelona: Empúries, 1994), 85-6.)

⁶¹ Stamper, *Word by Word: The Secret Life of Dictionaries*, 35-6.

that these words have been in wide use for an extended period. A language suited for all uses needs all usages to be adequately included in its dictionary. In this way, a colloquial standard will be able to appear, as per the recommendations of the Secció Filològica;⁶² and colloquial Catalan will be given the respect it is due in its most essential domain.

References

- Acadèmia Valenciana de la Llengua, *Diccionari normatiu valencià* (DNV) (València: Generalitat Valenciana, 2014), <avl.gva.es/lexicval/>, accessed 19 March 2019.
- Bassols, Margarida, ‘El col·loquial dels mitjans’, in Margarida Bassols, and Milà Segarra, eds, *El col·loquial dels mitjans de comunicació* (Vic: Eumo Editorial, 2009), 11–20.
- Bernal, Elisenda, and Sinner, Carsten, ‘Al seu rotllo: aproximació al llenguatge juvenil català’, *Zeitschrift für Katalanistik* 22 (2009), 7–36.
- Bernal, Elisenda, Costa, Joan, and Paloma, David, *La codificació i l'estandardització del català avui* (Barcelona: Fundació per a la Universitat Oberta de Catalunya, 2010).
- Bldas, Òscar, ‘La llengua oral col·loquial’, in Margarida Bassols, and Milà Segarra, eds, *El col·loquial dels mitjans de comunicació* (Vic: Eumo Editorial, 2009), 23–39.
- Cabré, M. Teresa, ‘De la descripció a la norma: el model de llengua de l’Institut d’Estudis Catalans’, in *Gramàtica de la llengua catalana* (Barcelona: Institut d’Estudis Catalans, 2016), XIII–XV.
- Corporació Catalana de Mitjans Audiovisuals, *ésAdir: portal lingüístic* (Barcelona: Generalitat de Catalunya, 2006), <esadir.cat>, accessed 21 June 2019.
- Diccionari manual de llengua catalana* (VOX) (3rd ed.) (Barcelona: Larousse, 2015).
- Fabra, Pompeu, *Diccionari general de la llengua catalana* (Barcelona: Edhasa, 1932).

⁶² Secció Filològica, ‘Sobre la variació i la prescripció en les orientacions normatives de la Secció Filològica’.

- Freixas, Martí, *Els sufixos derivatius apreciats en català: estudi morfo pragmàtic i lexicogràfic* (research plan) (Barcelona: Universitat Pompeu Fabra, 2018).
- García Márquez, Gabriel, 'La mujer que escribió un diccionario', *El País* (10 February 1981).
- Garriga, Cecilio, 'La marca de "vulgar" en el DRAE: de *Autoridades* a 1992', *Sintagma* 6 (1994), 5–13.
- Ginebra, Jordi, and Solà, Joan, *Pompeu Fabra: vida i obra* (Barcelona: Teide, 2007).
- González Pérez, Rosario, 'La marca *figurado* en los diccionarios de uso', *Revista de Lexicografía* 7 (2000), 77–90.
- Gran diccionari de la llengua catalana* (GDLC) (Barcelona: Enciclopèdia Catalana, 1998), <diccionari.cat>, accessed 10 March 2019.
- Institut d'Estudis Catalans, *Diccionari de la llengua catalana* (DIEC2) (2nd ed.) (Barcelona: Edicions 62; Enciclopèdia Catalana, 2007), <dlc.iec.cat>, accessed 14 March 2019.
- Institut d'Estudis Catalans, *Gramàtica de la llengua catalana* (GIEC) (Barcelona: Institut d'Estudis Catalans, 2016).
- López del Castillo, Lluís, *Llengua standard i nivells de llenguatge* (Barcelona: Laia, 1976).
- Matamala, Anna, 'Del guió a la pantalla: trets de l'oral col·loquial en la ficció televisiva', *Caplletra* 44 (2008), 141–168.
- Ortega, Rudolf, 'El fill bord del Termcat', *El País* (27 July 2015).
- Oxford English Dictionary, 'Have you ever wondered how words enter the *Oxford English Dictionary*' (Oxford: Oxford University Press, 2018), <<https://public.oed.com/how-words-enter-the-OED/graphic/>>, accessed 26 June 2019.
- Payrató, Lluís, 'Lexicografia i variació funcional', *Caplletra* 17 (1994), 47–58.
- Pla, Josep, *El carrer Estret* (Barcelona: Destino, 1951).
- Porto Dapena, José-Alvaro, *Manual de técnica lexicográfica* (Madrid: Arco Libros, 2002).

Salvanyà, Jaume, 'El tractament de la llengua col·loquial: de Fabra als nostres dies', in *Andorra i Pompeu Fabra. Recull de les conferències sobre Pompeu Fabra* (Andorra la Vella: Centre de la Cultura Catalana, 2019), 39–60.

Salvanyà, Jaume, 'Les marques de "baix" registre en els diccionaris', *Llengua Nacional* 72 (2010), 16–19.

Secció Filològica, 'Sobre la variació i la prescripció en les orientacions normatives de la Secció Filològica', in *Documents de la Secció Filològica, V* (Barcelona: Institut d'Estudis Catalans, 2018), 35–48.

Solà, Joan, *Sintaxi normativa: estat de la qüestió* (Barcelona: Empúries, 1994).

Stamper, Kory, *Word by Word: The Secret Life of Dictionaries* (New York: Pantheon Books, 2017).