

Received 6 January 2017.

Accepted 30 March 2017.

NOTES ON THE TREATMENT OF VARIATION IN THE *DICCIONARI CATALÀ-VALENCIÀ-BALEAR* (DCVB)¹

Isidor MARÍ

Institut d'Estudis Catalans*

imari@uoc.edu

Abstract

The article analyzes the variation tags used in the DCVB with the double purpose of facilitating their comprehension by the users and of suggesting future more precise characterizations. In the light of the recent criteria of variation, different works of Catalan, Spanish and international lexicography on the representation of the variation in the dictionaries are commented on, and a proposal of synthesis is presented as a conclusion.

Keywords

lexicography, variation tags or marks, Catalan

NOTAS SOBRE EL TRATAMIENTO DE LA VARIACIÓN EN EL *DICCIONARI CATALÀ-VALENCIÀ-BALEAR*

Resumen

El artículo analiza las etiquetas de variación usadas en el DCVB con la doble finalidad de facilitar su comprensión por parte de los usuarios y de sugerir futuras caracterizaciones más precisas. A la luz de los criterios recientes de variacionismo, se comentan diferentes trabajos de lexicografía catalana, española e internacional sobre la representación de la variación en los diccionarios, y se presenta como conclusión una propuesta de síntesis.

¹ This work has been developed under the FFI2013-41077-P project, funded by the Spanish Ministerio de Economía y Competitividad.

* Carrer del Carme, 47, 08001 Barcelona.

Palabras clave

lexicografia, etiquetas o marcas de variación, lengua catalana

Feci quod potui, faciant meliora potentes.

Francesc de B. Moll, at the end of the *Introducció* to the DCVB

1. Introduction

We do not have the pretension, with these simple notes, to improve the excellent work by Alcover and Moll — and all the other collaborators who intervened and who have allowed us to have this great lexicographic work today. Our purpose is much more modest: to explore which treatment (linguistic) variation receives in the DCVB, with two primary purposes. Firstly, and above all, in order to make it easier for users of this work to understand the orientations that can be found about the use and value of each of the lexical units; and secondly, in order to consider the feasibility of a future complementary work of more detailed characterization of the variation in the *Diccionari català-valencià-balear* (DCVB).

Obviously, the fact that in recent years different sorts of variation — diachronic, diatopic, diastratic and diaphasic, according to E. Coseriu (1981) — must be included in lexicographic works, according to their type, we cannot forget that at the time of the DCVB's publication many of these considerations were unusual, as will be discussed below.

2. The marks of variation in contemporary Catalan lexicography of the DCVB

Lluís Payrató (1994), in one of the first works on the study of the variation in the Catalan lexicography fieldwork, already underlined that even in the *Diccionari general de la llengua catalana* (DGLC) by Fabra (1932), which was a primordial reference

during the elaboration of the DCVB, the marks of variation are limited and not entirely systematic:

As for the historical varieties, the *antigament* 'formerly', *arcaic* 'archaic' and *obsolet* 'obsolete' expressions coexist, and as for the functional varieties, *propí del llenguatge familiar* 'typical of the familiar language' and *popularment* 'popularly' (and still, in part, *habitualment* 'habitually' / *ordinàriament* 'ordinarily', on the one hand, and *emprat irònicament* 'ironically used' / *emprat poèticament* 'poetically used', on the other). If we add *vulgarment* 'vulgarly', which has in principle to do with sociocultural variation, and the very long series of expressions referring to the technical field, the general view is complete with respect to the DGLC.

Payrató's work focuses on the analysis of the functional variation, but it has the advantage — with respect to other similar studies — of starting from a consistent sociolinguistic basis on variationism, which allows him to establish a fairly clear identification of the different types of variation:

- a) the variation that corresponds to geographical, social, or chronological dialectal varieties (historical or generational);
- b) the variation related to functional varieties or registers, which in turn depend on determined factors (which are shown in parentheses):
 - The thematic field (general / specialized)
 - The mode or channel of communication (oral / written, spontaneous / planned)
 - The interpersonal tenor (formality / familiarity)
 - The purpose or enunciative attitude (ironic, humorous, contemptuous, poetic)

This sort of systematic approach is not found even in the majority of lexicographical works after Fabra: the factors of functional variation are rarely taken into account, and the labels or marks used therein are not usually accompanied by precise definitions that account for the criteria with which they have been established and applied. In fact, even today, although scholars talk of the need to incorporate

sociolinguistic criteria in lexicographical works — and some authors even talk of socio-lexicographic approaches —, the need to systematize the variation labels is still a discussed matter and it lacks sufficiently general, accurate and clear solutions. First, because it is not easy to establish a valid system of variation marks for all types of lexicographical works — monolingual or bilingual, descriptive or prescriptive, dialectal or centred in the standard variety, etc.; and above all, because it is essential to obtain a satisfactory and considered commitment between the aim of reflecting the variation with the utmost precision and detail and the need to provide users with transparent criteria for data interpretation and also useful for their communicative practice.

According to Payrató, among the factors of functional variation — or the register —, the thematic field is the one that is more usually reflected in detail in the lexicographic works — also, as we shall see, in the DCVB — with marks and distinctions of the different specialized fields of knowledge. The mode or channel of communication does not appear much, even in the basic distinctions between lexical units specific of spoken language — or of spontaneous use as opposed to planned. The interpersonal tenor (or degree of formality) usually is not reflected more than in some inaccurate labels, as *familiar* ‘familiar’, *popular* ‘popular’ or *vulgar* ‘vulgar’. And the tone or purpose of the communication does not usually go beyond distinctions such as *irònic* ‘ironic’, *despectiu* ‘derogatory’, *figurat* ‘figurative’ or ‘o poètic’ ‘poetic’.

3. Variation marks in other lexicographic contexts

The works that deal with the marks of variation in Spanish lexicography (for example, Garriga 1997; Fajardo 1996-1997) also coincide in the lack of theoretical justification and the notable fluctuation that the different lexicographic works present in the use of variation marks.²

² More recently, López Morales (2010) also comments on the state of the issue, and considers that dialectal dictionaries should provide sociolinguistic information of four types: a) the specific register to which the word belongs, b) its relation to a certain sociocultural level within the sociolinguistic spectrum of the studied community, c) the social value that the community itself makes of the word, and d) the indication of the language style to which the word belongs. Slogans or meanings that do not carry any of these marks should be interpreted as socially neutral. However, the examples he mentions in order to illustrate his proposal are not particularly convincing.

Garriga's study is eminently descriptive, but Fajardo's aim is to critically examine the marks and deserves to be given more attention.³

Fajardo rightly starts from the need to clarify the concept of marking (*marcación*) as a basis for identifying which lexical units are marked or unmarked, according to particularities that condition or restrict their use.⁴ It is also relevant his observation on the difference between dictionary abbreviations and the marks themselves: not all abbreviations are marks, not all marks are manifested by means of abbreviations.

Either way, Fajardo also emphasizes that a satisfactory systematization of use marks in Spanish language dictionaries has not been achieved: neither the name nor label of the marks coincides, nor does the value of the same marks. However, within this heterogeneity and the relative value they have, we must recognize the need and utility of the marks.

Another interesting point of Fajardo's article is the typology that it presents of the oppositions reflected in the marks and of the labels with which they are represented in the dictionaries:

- 1) actual 'current' // nuevo 'new' / obsolescente 'obsolescent' / obsoleto 'obsolet': *neol., desus., ant.*
- 2) general // regional / dialectal: *Amér., And., Argent., Al.*
- 3) neutro 'neutral' // hablado 'spoken' / escrito 'written': *lit.*
- 4) neutro 'neutral' // estrato 'stratum' / grupo 'group': *po., vulg.*
- 5) neutro 'neutral' // formal / informal: *coloq., form., inform.*
- 6) neutro 'neutral' // poético 'poetic' / literario 'literary' / periodístico 'journalistic' / administrativo 'administrative': *Adm., lit., Poét.*
- 7) lengua común 'common language' // lengua técnica 'technical language': *Acúst., Aer., Agri., Agrim.*
- 8) frecuente 'frequent' // infrecuente 'infrequent': *desus., us.*
- 9) neutro 'neutral' // connotado 'connoted': *despect., irón., insul.*

³ On the other hand, it is the article to which the RAE dictionary refers for the understanding of the marks that it uses (see [drae.es/marcas](http://d.rae.es/marcas) [accessed in January 2017]).

⁴ Perhaps it would be more accurate to say that the marked units are accompanied by connotations that the user must have in mind, although they do not condition or restrict their use inflexibly: after all, it is imperative that the user assumes the availability of all the lexical units.

10) correcto 'correct' // incorrecto 'incorrect': *Incor.*, *.

Some of these oppositions, according to the author, are antonymic (such as 10); others respond to a gradation or continuum (such as 8), and others have an unmarked pole and a multiplicity of marked features, such as 7. However, Fajardo does not refer to the explanatory factors of these oppositions, and although some of them clearly correspond to a specific factor of variation (such as 10, which refers to the normativity of lexical units), some oppositions are not very precise: (1) and (8) overlap, and apparently (4), (5), (6) and (9) also have to be delimited better.

In short, the conclusions reached by Fajardo admit that the complexity of all the factors of variation makes it difficult to reach a general and complete systematization:

- The use of marks is frequent in Spanish lexicography, but they have not been normalized.
- Marks should be reviewed and unnecessary or repetitive marks should be avoided.
- The selected marks should be clearly and easily graded.
- The index of abbreviations should provide information on the value of each mark.
- It is worth making an effort to systematize the marks because they are useful for users.

4. The lexicographical treatment of variation, a complex question

A brief review of some recent works on the treatment of variation in lexicographical works, in different linguistic and cultural contexts as well as in different types of dictionaries, confirms the difficulty of finding a fully systematic and fairly precise solution, and at the same time sufficiently transparent and clear for users.

According to Beyer (2011), although marks or variation tags are apparently similar among the lexicographic works, there are significant differences in the deeper classifications and subclassifications, and a general consensus about the use of lexicographical marks and the pragmatic parameters which they represent is absent,

due to the lack of a common and standardized theoretical basis. The clarification of these parameters and labels depends on the concept of *marked* and its specific application in the lexicographic field, which does not exactly coincide with the linguistic notion of *marked*. On the other hand, the criteria vary according to the type of dictionary and the audience to which it is addressed.

Based on these criteria, Beyer proposes a typology of lexicographic labels, which aspires to be complete, without exhaustiveness:⁵

1. **Domain labels**, which indicate the assignment of lexical units to a given domain.

a. *Geographic labels*, which indicate the spatial distribution.

b. *Temporary labels*, which indicate the chronologically marked character of the units.⁶

i. *Period labels*, which indicate the chronological position with respect to the contemporary moment (e.g., *antiquat* 'old-fashioned' o *neologisme* 'neologism').

ii. *Historical labels*, which indicate items that are no longer in use.

c. *Frequency labels*, which indicate a marked frequency of use.

i. Absolute *frequency labels*, such as *rar* 'rare'.

ii. *Relative frequency labels*, such as *poc usual* 'unusual'.

d. *Technical labels*, which indicate the ascription to a specific field of knowledge.

e. *Cultural labels*, which indicate the ascription to a certain area or cultural tradition.

2. **Linguistic labels**, which indicate the linguistic exceptionality of the units depending on the type of dictionary.⁷

3. **Stylistic labels**, which indicate the stylistically marked character of the units.

a. *Register labels*, which indicate the ascription to specific situations of use, such as *periodística* 'journalistic' or *poètic* 'poetic'.

⁵ The author emphasizes that each dictionary must choose its own system of marks and labels, adjusted to the needs of the users to whom it is addressed.

⁶ Chronological marks are especially important in dictionaries that focus on the contemporary language. These types of labels and also frequency labels can be integrated.

⁷ We understand that this section would include the marks and labels referring to the prescriptive or standard character.

b. *Socio-stylistic labels*, which indicate the viability of the lexical units in a certain level or style, such as *formal* or *vulgar*.

Another general approximation to the lexicographical marks of variation that I find particularly interesting is that of Marek Maziarz *et al.* (2015), of the Polish team of WordNet, oriented to the computer processing of the lexicon.

In their approach, the criterion is that there are different scales of variation, each of which has a neutral or unmarked area, corresponding to the general variety:

- The **emotional** scale, between *appreciative (kind)* and *offensive*.
- The scale of **formality**, between *high* and *intimate*.
- **Frequency** of use, between very *frequent* and *unusual*.
- The scale of **genuineness**, between *loanword* and *original*.
- The scale of **textuality**, between *poetic* and *colloquial*
- The **diatopic** scale (or continuum), between *local* and *general* (or international).
- The scale **diastratic**, between *neutral* and *argot*.
- The **dianormative** scale, between *correct* and *non-standard*.

These authors have proposed to simplify and hierarchize these marks or labels, and have reduced them to eleven, which would be applied following a decision tree, which they represent in a scheme (Figure 1). Their proposal is attractive, but in our opinion it is not without drawbacks, especially in the final phase of the model.

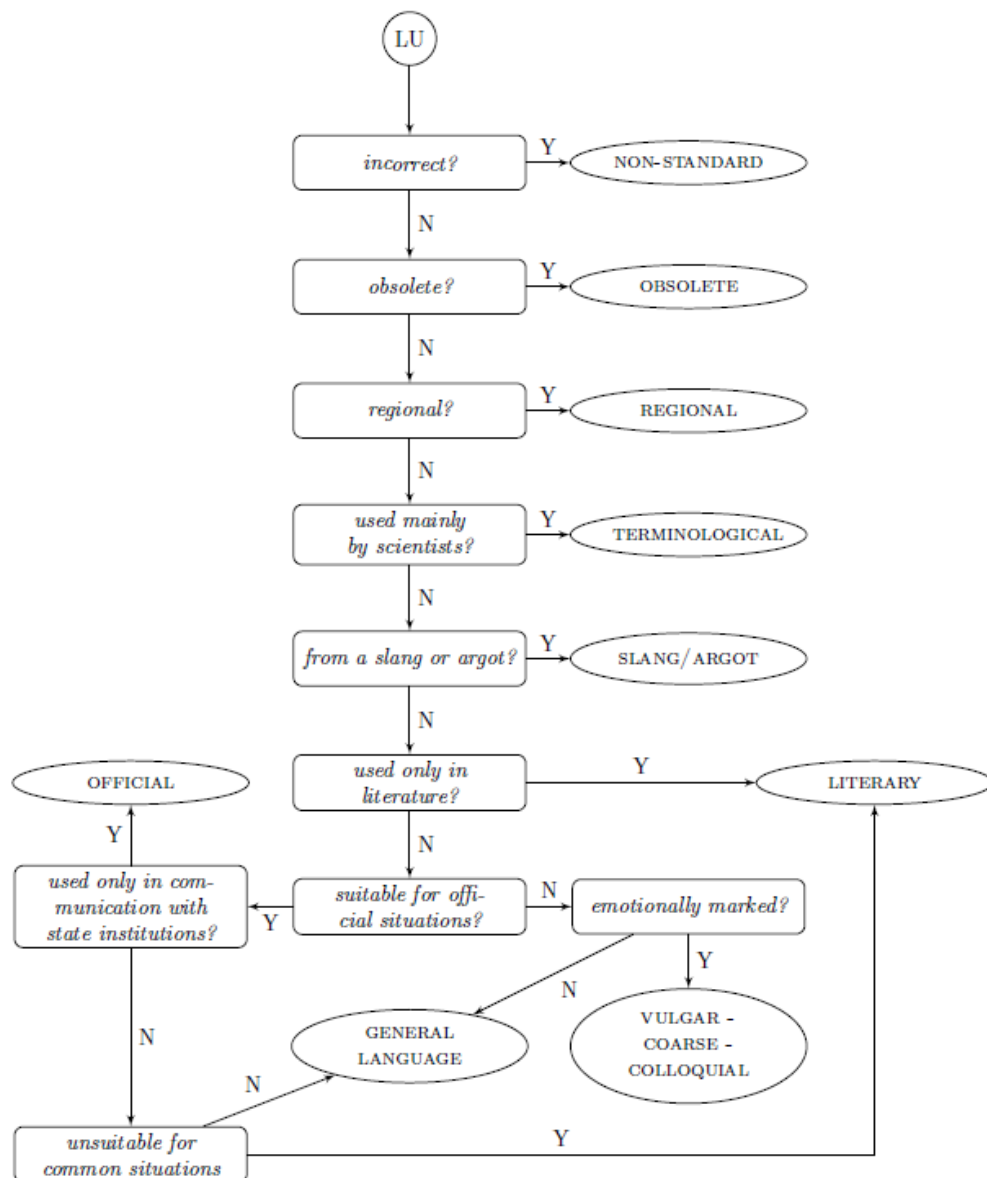


Figure 1: The decision tree for register assignment. The tests for three emotive labels have been conflated.

5. Variation marks in the DCVB

The treatment of variation in the DCVB is explicitly explained in the *Introduction* of the work, specifically in the second part, *Description of the dictionary*.⁸ From the first paragraph it is stated that “*es tracta d’un diccionari integral, que reuneix els*

⁸ P. XXII ff of the first volume (according to the second edition, dated 1975).

materials de la llengua literària, dialectal, antiga i moderna".⁹ As usual, the *literary language* is understood as the common written or supraregional variety, whereas the *dialectal language* makes reference to the usual talks of the different places of the linguistic territory.

A first section of the introduction is dedicated, in fact, to *the spoken language*, that is to say to the "*llenguatge viu, col·loquial, no literari*" 'living, colloquial, non-literary language'. The second, to *the written language* — specifying the criteria for using selected works as sources for obtaining words. This second section does not contain relevant information regarding the treatment of historical variation. A third section explains the *Structure of the lexical entries*.

6. The spoken language: geolectal variation

In the case of the spoken language, lexical units are linked to the localities in which each has been documented, although the means of collecting them are quite diverse: surveys with different questionnaires, addition of a questionnaire of "*totes les particularitats de lèxic i de gramàtica que sortien espontàniament en la conversa dels individus interrogats o de qualssevol altres habitants de la localitat*",¹⁰ contribution of numerous collaborators who "*donaven preferència, com és natural, als mots que consideraven típics de llur vila o comarca*".¹¹ Therefore, the introduction includes "*quan indicam per a un mot –o per a un significat especial d'un mot– els llocs on ha estat recollit, oferim almenys una orientació molt útil per a aclarir el problema de les àrees lèxiques*".¹²

An attached map shows the localities in which Alcover carried out the survey alone, together by Alcover and Moll or by Moll and Sanchis Guarner individually. At the

⁹ "... it is a comprehensive dictionary, which brings together the materials of the literary, dialectal, ancient and modern language".

¹⁰ "... all the characteristics of lexicon and grammar that came spontaneously in the talk of the informants questioned or of any other inhabitants of the locality".

¹¹ "... who, of course, gave preference to the words that considered typical of their city or region".

¹² "... when we ask for a word or for a special meaning of a word the localities where it has been collected, we offer at least one very useful orientation to clarify the problem of the lexical areas".

end of the section devoted to the spoken language the localities in which there were especially active collaborators are also detailed.

The characterization of the geographic variation is completed later in the first part of the *Explanation of the abbreviations*, devoted to the *Geographical abbreviations*. In this case, next to the names of specific localities, we find references to dialect areas: *alg.* (Alguerès dialect), *bal.* (Balearic dialect), *cat.* (Catalan), *eiv.* (Ibiza dialect), *mall.* (Majorcan), *men.* (Menorcan), *occ.* (western Catalan dialect), *or.* (eastern Catalan dialect), *pir.-or.* (Eastern Pyrenean dialect or Rossellonès), *val.* (Valencian), and also references to more or less extensive geographical areas *-bal.* (Balearic Islands), *Cat.* (Catalonia), *Maestrtr.* (Maestrat), *Mall.* (Majorca), *Men.* (Menorca), *Rib. d'Ebre* (Ribera d'Ebre), *Ross.* (Roussillon), etc. The dictionary query suggests that there are more references to other localities that are not listed on the map or in the list of abbreviations.

Additionally, in the section devoted to the *Explanation of the abbreviations* we can find a generic mark, apart from those mentioned in the previous paragraph: *dial.* (dialectal). In these cases, it seems that the mark is intended to make the difference with respect to the general form adopted by the prescriptive lexicon.¹³

7. The structure of the lexical entries: other marks of variation

The explanation of the ordering criteria of the lexical entries details how the variation is presented, in which the *documentary part* of each entry is called: thus, in parentheses, it is stated “*les **regions**¹⁴ o **pobles** on hem recollit aquella accepció, **si no és d'ús general a tot el territori**”¹⁵.*

An implicit way to characterize the most general form of a lexical unit is to refer from the most local variants to the main entry, headed by the most general form or at

¹³ See the example discussed below, from the lexical entry *abantes* ‘before’.

¹⁴ The bold is ours: note that the lexical units of general use are not marked.

¹⁵ “... the **regions** or **localities** where we have collected that meaning, **if not it is of general use throughout the territory**”.

least presenting this first. This is done, for example, at the lexical entry of **batejar** or **batiar** ‘to baptize’, to which the *batiar* and *bateiar* forms refer.

It is interesting to note that the section on phonetics (FON.) includes the phonetic transcription of the various regions in which the lexical unit has been documented, with the greatest detail: *“Hem procedit amb gran esment en la transcripció de les paraules d’ús popular o dialectal, marcant totes les diferències que hem pogut observar en la pronúncia viva, fins i tot les que són incorrectes, perquè totes tenen importància per al dialectòleg i per a l’historiador de l’idioma. Les paraules cultes (neologismes, llatinismes, hel·lenismes) les transcrivim només segons la pronúncia pròpia de les tres capitals Barcelona, València i Palma de Mallorca”*.¹⁶

From a point of view of historical or chronological variation, the information provided in the section on *spelling variants* (VAR. ORT.) is interesting, *“on es registren les grafies arcaiques que convé fer constar per a documentació dels estudiosos de la història de la llengua.”*¹⁷

From an even a more general point of view, the paragraph *variants of formation* (VAR. FORM.) *“presenta les diverses formes, antigues o dialectals, del mot cap d’article”*.¹⁸ In this case, the relationship can be more formal than meaning, as in the case of the example adduced therein, which relates to *balda* and *baula* ‘latch’.

8. Other abbreviations

In addition to the cases mentioned above, relating mainly to geolectal variation, the abbreviations of the DCVB contain several references to other types of variation:

-The adscription to a thematic field or a field of knowledge: *acúst.* (acoustics), *agric.* (agriculture), etc.

¹⁶ “We have proceeded with great care in the transcription of words of popular or dialectal use, underlining all the differences that we have been able to observe in the living pronunciation, even those which are incorrect, because they all are important to the dialectologist and to the historian of the language. We transcribe the educated words (neologisms, Latinisms, Hellenisms) according to the pronunciation of the three capitals Barcelona, Valencia and Palma de Mallorca”. Note also this contrast between the lexical units of **casual or dialectal use** and the **educated words**.

¹⁷ “... where archaic spellings are registered, which should be recorded for the documentation of scholars of the history of language.”

¹⁸ “... presents the various forms, old or dialect, of the word head of lexical entry”.

-The adscription to another language or variety: *alem.* (German), *alt-alem.* (High German), *angl.* (English), *aràb.* (Arabic), *b.-llatí* (low-Latin), *cast.* (Castilian), *esp.* (Spanish), etc.

- Two marks of chronological variation: *ant.* (old)¹⁹ and *neol.* (neologism).²⁰

- Two more marks of figurative meaning (*fig.*) or metaphorical (*met.*).²¹

- Some marks of pragmatic value, such as *eufem.* (euphemism) and *pejor.* (pejorative),²² and stylistic: *poet.* (poetic).²³

- The abbreviations *pop.* (casual, popular), *vg.* (vulgar) and *vulgar.* (vulgar, vulgarism).²⁴

Apart from the abbreviations included in the initial list, occasionally we can find others, such as *iron.* (Ironically), to the entry *afaitar* 'to shave', next to the meaning of "to deceive". On the other hand, there are cases in which the figurative or ironic value is not mentioned: *albercoc* 'apricot', with the sense of "Bajà, poc-seny, irreflexiu" 'Silly, foolish, thoughtless'.

9. Discursive references to the variation in the body of the lexical entries

A brief review of some lexical entries allows us to verify that much information about variation is included, discursively exposed. Some examples demonstrate this fact.

¹⁹ This indication often appears next to unusual units or spellings (eg *abcegar* 'to blind' or *abbadessa* 'abbess'). In some cases, as *abciac* 'fateful', it refers to the spelling "modern adopted", *atziac*. In the entry *adalil* 'warlord' in addition to the abbreviation *ant.* 'old' explains: "it is an archaic word, re-established in the written language". It should be understood that the dating of the written sources also provides some chronological information about the lexical units.

²⁰ In this case, it usually only appears in terminological units, such as *abioquímica* 'abiochemistry' and *abissodinàmica* 'abissodynamics'.

²¹ Thus, the entry *abordar* 'to board' has a meaning which includes the *met.* mark: *Renyar fort* 'strong rebuke'.

²² This abbreviation is commonly used in certain derivatives.

²³ The entry *acer* 'steel', for example, includes the mark *poet.* "Arma d'acer (espasa, punyal, etc.)" 'Steel weapon (sword, dagger, etc.)'.

²⁴ However, we will see later that this notion does not seem to have rude or malicious social connotations, but refers to popular expressions that are not linguistically correct. This is the case, for example, of one of the meanings of *abusió* 'abuse', marked with the abbreviation *vg.* (vulgar), when used improperly with the sense of "Gran multitud, excés" 'great multitude, excess'.

abans ‘before’

- It includes discursive references to formal variants, such as:

Sometimes it is preceded by the prep. *de*. “Jo, d’abans hi anava força, a l’església” ‘Before I used to go to church a lot’ (Gir. ‘Gerona’)

- Next to the phonetic transcription there are some indications on the extent of use:

In many regions it disappears from the living language, yielding its place to *antes* ‘before’.²⁵

- In the SINÒN. (synonym) and VAR. FORM. (variants of formation) sections refers to *ans*, *antes*, *abantes*, *abant*.

The consultation of these other forms gives us additional information, especially etymological and historical. However, the lexical entry **antes** ‘before’ marks this form as *castilianism*²⁶ and in the section SINÒN. considers ‘preferable’ *abans* and *ans*. We also note that the entry **abantes** includes the mark *dial.* (dialect) which probably informs not only of the territorial distribution of this lexical unit, but also that it is a form not admitted by the prescriptive dictionary.

In some cases it is surprising that there is no indication of the familiar, popular, informal or even vulgar or malicious character of certain lexical units. For example, the series of expressive forms related to **amagar** ‘to hide’ — (d’) *amagatejons*, (d’) *amagatons*, *amagatontes*, (d’) *amagatotés*, (d’) *amagatotis* — do not carry any

²⁵ In the entry *abraçadora* ‘clamp’, for example, this can be read: “Està molt estesa la forma *abraçadera*, presa directament del cast. *abrazadera*”. “The word *abraçadera* ‘clamp’ is widespread, taken directly from the Cast. *abrazadera*. In other cases the information about the loss of vitality of a lexical unit is also given discursively. This is the case of the word *abadí* ‘marble’, for example: “El mot *abadí*, no l’hem sentit mai en el llenguatge viu de Mallorca; només el coneixem pel que en diuen els dos diccionaris citats [*Amengual* and *Un Mall. Dicc.*]” ‘We have not heard the word *abadí* in the living language of Majorca; we only know it from what is written in the two mentioned dictionaries’.

²⁶ This mark often appears next to the lexical entries. For example, it is also included in the entries *abadeco* ‘cod’ and *abarcar* ‘embrace’. However, the DCVB does not always reject the use of these types of words. In the case of *abarcar*, the DCVB proposes in the SINÒN. section that the words *abraçar*, *agafar* are preferable. On the other hand, in the lexical entry *abono* ‘fertilizer’, it indicates that the *abonament* “és el mot que cal emprar en lloc de *abono* (castellanisme)” ‘it is the word we must use instead of *abono* (Castilianism)’. This happens also in the word *acomodo* ‘arrangement’: “Cal evitar l’ús d’aquesta paraula en català, substituint-la amb sos sinònims *acomodació* o *acomodament*” ‘We must avoid using *abono* in Catalan, replacing it with its synonyms *acomodació* and *acomodament*’.

indicative of functional or pragmatic characterization, although all of them refer to *d'amagat* 'on the quiet, secretly'.

A similar situation occurs (with apologies) in the case of the exuberant family related to **cagar** 'to shit', which is not marked as vulgar.²⁷ Only in some cases does it receive a tangential reference because of the connotations that accompany it:

Cagar-se en algú o en alguna cosa 'To hell with someone or something': *menysprear-la* insolentment 'defiantly despise it'; cast. *cagarse*. This meaning is the basis of blasphemies and profanity ("me cago en..."), which sometimes it is softened by euphemism by changing *cago* 'to hell' in *caso* 'I get married' ("me caso en...") or replacing the sacred name object of the blasphemy by a harmless word ("me cago en la mar salada", "me cago en dena", "me cago en Sivilla", etc.).

Or in the entries *caganiu* and *caganius*, which allude to the contemptuous tone that occasionally has this word: "*pl. Habitants de Darnius. És un nom despectiu que els donen els nadius dels pobles veïns.*" 'Darnius inhabitants. It is a derogatory name given to them by the natives of neighbouring villages'.

Neither do we find in these entries the mark of figurative meaning (*fig.*) in cases in which would seem logical, as the meaning of *cagada* 'what a fuck up!' as *Acció ridícula o d'un èxit infeliç, contrari al que s'esperava obtenir* 'ridiculous action or unhappy success, contrary to what was expected to be obtained'. Or also *cagalló's* meanings such as "*Home covard (Barc.)*" 'cowardly man (Barc.)', "*Home despreciable, que no té caràcter (Mall., Men.)*" 'despicable man, who has no character (Mall., Men.)' or even "*Infant mal criat (Mall., Men.)*" 'brat' (Mall., Men.).²⁸

Therefore, the distinction of degrees of formality (or of the intentionalities of the discourse) does not seem very present in the characterization of the lexical units of the DCVB. However, we can find some discursive reference to the stylistic value of some words, such as in the entry *absort* 'absorbed', which includes "*És paraula d'ús purament literari*" 'This word has a purely literary use'.

²⁷ However, the value of *vulgar* is associated with pronunciations such as *ausència* 'absence' and *ausent* 'absent', as opposed to those described as literary, *absència* and *absent*, respectively. It must be understood that it is a linguistic concept, unrelated to the social connotation of rude or malicious.

²⁸ Occasionally the pragmatic value of some lexical unit is explained discursively, as in the case of *abraguerat*: "*Ho solen dir irònicament, de qualcú a qui ha succeït qualche cas advers*" 'They usually say it ironically from someone who has had an adverse case'.

10. In conclusion

If we return to the well-founded conceptual distinctions by Payrató (1994) and confront them with the attempts to systematize other aforementioned works with a view to a prospective completion of the DCVB variation marks, we could formulate the following proposals:

a) Geolectal variation: general units (unmarked) / geographic location markings (local, regional, or generically dialect).

These are currently the most detailed marks in the DCVB. Probably a review would allow us to homogenize and explain those which are not now either on the map or in the list of abbreviations. It is also necessary to check the words currently marked as a *dial.* (dialectal) that must adopt the mark of prescriptively not accepted.

b) Diachronic variation: old units (unused) / units available in contemporary use (obsolete / unmarked / recent / neological). In addition, the inclusion of relative frequency indicators (rare / unusual / current: unmarked) can be evaluated.

There are lexical units, spelling variants and variants of formations that already include the mark *ant.* (old) in the DCVB. A review would make it possible to check whether other words should also be marked as outdated (or obsolete, if they are still used now improperly). Similarly, it should be evaluated whether terminological units currently marked with the *neol.* mark must be kept it. We have also seen that in some cases there are entries of the DCVB that contain discursive information about the recent extension of use of some elements: this appraisal corresponds to the moment of the 20th century when the DCVB was written and it would be advisable to consider if it is still logical to be marked as recent. As regards frequency marks, if it is considered convenient to include them, the frequency indication of the *Diccionari Descriptiu de la Llengua Catalana* (DDLCC) of the Institut d'Estudis Catalans might be taken as reference.

c) Sociolectal variation: units not marked / ascribed to a certain social group / jargon.

These sorts of marks are rare in the DCVB and on the other hand we must take into account that sociolects are difficult to identify in today's society. In any case, it is convenient not to confuse these marks, which must be reserved to specific features of

social groups, with those corresponding to register factors linked to communicative situations (see sections f, g and h).

d) Genuine or normative marks: unmarked (genuine and accepted by the prescriptive grammar) / not allowed (with possible indication of the language of origin).

We have seen that the DCVB sometimes uses the mark *dial.* (dialectal) as an indication that the lexical unit has not been accepted by the prescriptive dictionary. Other times it specifies that they are Castilianisms. There may be other cases of borrowing or interference not prescriptively accepted. A review would make it possible to unify these marks and to decide whether it would be useful to specify the source language of cases not formally accepted or not adapted. Again the DDLC can be a reference for the concretion of the mark of not prescriptively accepted.

e) Thematic field marks: unmarked (current language) / mark of the respective field of knowledge.

As we have mentioned before, these marks are very abundant in the DCVB. A revision – perhaps with the advice of the TERMCAT (the Centre for Catalan Terminology) – would allow for the systematization of the assignment of the words to an updated thematic index.

f) Formality marks: unmarked / formal (possibly very formal) or informal (familiar, childish, popular, vulgar).

Often these types of marks — which must be linked to the degree of formality of communicative situations — are mistaken for those of intentionality or emotional tone of the communication (see section h). Thus, it is necessary to assess which lexical units must be linked to markedly formal or solemn situations (in which the distance or lack of familiarity between the partners is clear) and which, on the contrary, are indicative of informal situations — which in some cases can be marked more specifically as childish (child-specific) or vulgar (not in the sense of improper formal variants now used in the DCVB, but as socially connotated as coarse or rude). Familiar or popular marks do not seem necessary.

g) Mode or channel marks: unmarked / specific to a particular mode of communication (oral / written, colloquial / literary or poetic).

We have seen that some lexical entries in the DCVB use the mark *poèt.* (poetic). In other cases the merely literary character of a word, usually of the ancient language, is commented discursively. One could consider the inclusion of all these cases in a single mark of literary connotation. There may be lexical units that are markedly specific to oral language or of certain writing formulas. The colloquial mark should be reserved for units or features (formal or phonetic) of spontaneous speech, avoiding possible confusion with the mark of informality.

h) Marks of intentionality or emotion: ironic, derogatory or pejorative, figurative, euphemistic.

There are marks that already appear in the DCVB as ironic, derogatory or euphemistic (sometimes explained discursively). We also occasionally find the mark *met.* (metaphorically), which could be unified with *fig.* (figurative sense). The abbreviation *pejor.* (pejorative) seems to refer to the nominal derivative suffixes. It would be convenient to unify the current discursive indications in the form of an abbreviation and to evaluate the desirability of marking lexical units that now have no information.

We hope that these notes may be useful if at any moment a systematization of the variation marks of DCVB is feasible. It would be desirable that these marks coincide with the system of markers of the main Catalan lexicographic repertoires, especially the *Diccionari Descriptiu de la Llengua Catalana* (DDL) and the *Diccionari de l'Institut d'Estudis Catalans* (DIEC).

References

- BEYER, Herman L. (2011) '*n Algemene tipologie van leksikografi ese etikette* (Una tipologia general d'etiquetes lexicogràfiques), *Tydskrif vir Geesteswetenskappe*, year 51, n. 3, September 2011.
- COSERIU, Eugenio (1981) "La socio- y la etnolingüística", *Anuario de Letras. Lingüística y Filología*, 19, 5-30.
- DCVB = ALCOVER, Antoni M. & Francesc de B. MOLL (1975)² *Diccionari català-valencià-balear*, Palma: Ed. Moll.

DDLc = *Diccionari descriptiu de la llengua catalana*, Institut d'Estudis Catalans
<<http://dcc.iec.cat/ddlc/index.asp>>.

DIEC = *Diccionari de la llengua catalana*, Institut d'Estudis Catalans <<https://dlc.iec.cat/>>

DGLC = FABRA, Pompeu (1932) *Diccionari general de la llengua catalana*, Barcelona: Llibreria Catalònia.

FAJARDO, Alejandro (1996-1997) "Las marcas lexicográficas: concepto y aplicación práctica en la Lexicografía española", *Revista de Lexicografía*, III, 31-57.

GARRIGA, Cecilio (1997) "Las 'marcas de uso' en los diccionarios del español", *Revista de Investigación Lingüística*, 1, 75-110.

LÓPEZ MORALES, Humberto (2010) "La marcación sociolingüística en la lexicografía dialectal", in R. M^a Castañer & V. Lagüéns (coord.), *De moneda nunca usada. Estudios dedicados a José M^a Enguita Utrilla*, Zaragoza: Instituto Fernando el Católico, 385-391.

MAZIARZ, Marek, Maciej PIASECKI & Stan SZPAKOWICZ (2015) "The system of register labels in plWordNet", *Cognitive Studies / Études cognitives*, 15, 161-175.

PAYRATÓ, Lluís (1994) "Lexicografia i variació funcional", *Caplletra*, 17, 47-58.