Rute Costa, Margarida Ramos, Ana Salgado, Sara Carvalho,
Bruno Almeida, Raquel Silva

Neoterm or neologism? A closer look at the determinologisation process

1 Introduction

This paper arises within the current communication urgency experienced throughout the pandemic. From its onset, several new lexical units have permeated the overall media discourse, as well as social media and other channels. These units convey information to the public regarding the 'severe acute respiratory syndrome' namely *COVID-19.*[1] In addition to its worldwide impact healthwise, the pandemic generates noteworthy influence in the linguistic landscape, and as a result, a significant number of neologisms have emerged. Within the scope of our ongoing research, we identify the neologisms in European Portuguese that are related to the term *COVID-19* via form or meaning. However, not all the new lexical units identified in our corpus containing *COVID-19* in its formation can unequivocally be regarded as neoterms (terminological neologisms). Accordingly, this article aims not only to reflect on the distinction between neologism and neoterm but also to explore the determinologisation process that several of these new lexical units experience.

Following the introduction, this paper is divided into 9 sections. In section 2, we begin by making a brief theoretical reflection concerning neological processes

---

[1] In this paper, the term *COVID-19* is the preferred form referring to 'acute respiratory syndrome'.

Rute Costa, Margarida Ramos, NOVA CLUNL, Centro de Linguística da Universidade NOVA de Lisboa, Campus de Campolide, Colégio Almada Negreiros, 1099-032 Lisboa, Portugal

Ana Salgado, NOVA CLUNL, Centro de Linguística da Universidade NOVA de Lisboa, Campus de Campolide, Colégio Almada Negreiros, 1099-032 Lisboa; ACL, Academia das Ciências de Lisboa, Instituto de Lexicologia e Lexicografia da Língua Portuguesa, Rua da Academia das Ciências de Lisboa, 19, 1249-122 Lisboa, Portugal

Sara Carvalho, CLLC, Centro de Línguas, Literaturas e Culturas da Universidade de Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal; NOVA CLUNL, Centro de Linguística da Universidade NOVA de Lisboa, Campus de Campolide, Colégio Almada Negreiros, 1099-032 Lisboa, Portugal

Bruno Almeida, ROSSIO Infrastructure, Colégio Almada Negreiros, Campus de Campolide da NOVA, 1099-085 Lisboa, Portugal; NOVA CLUNL, Centro de Linguística da Universidade NOVA de Lisboa, Campus de Campolide, Colégio Almada Negreiros, 1099-032 Lisboa, Portugal

Raquel Silva, VOH.CoLAB – Value for Health CoLAB, Nova Medical School, Edifício CEDOC I, Rua do Instituto Bacteriológico, n.os 5, 5-A e 5-B, 1150-190 Lisboa, Portugal; NOVA CLUNL, Centro de Linguística da Universidade NOVA de Lisboa, Campus de Campolide, Colégio Almada Negreiros, 1099-032 Lisboa, Portugal

and determinologisation. Then, in section 3, we describe the method used to compile the corpus, the *CoronaCorpus*, which unfolds in 2 sub-corpora: the Press*CoronaCorpus* and the LSP*CoronaCorpus*. The Press*CoronaCorpus* is composed of texts published in the Portuguese media between November 2020 and July 2021. This corpus has been processed via Sketch Engine,[2] with the purpose of identifying neological lexical constructions occurring in non-specialised communication related to the emergence of the pandemic. Among such neological constructions, both neologisms and neoterms were identified. The latter are defined as terms that are "specifically coined for a given general concept" (ISO 1087:2019, §3.4.12). The second sub-corpus, the LSP*CoronaCorpus,* is composed of official documents produced by healthcare agencies, professionals and scientists. In the context of this research, this corpus plays the role of a reference corpus.

In section 4, our corpus is explored by means of simple and advanced queries for extracting the spelling variants of *COVID-19*. Section 5, on the other hand, is focussed on the *COVID-19* acronym, its behaviour in discourse and the re-categorisation of *covid-* as a formative in Portuguese. In section 6, we then proceed with the analysis of morphosyntactic and semantic formation of the neologisms and neoterms identified in the Press*CoronaCorpus,* to better grasp the process underpinning neology, in what concerns both form and meaning. Furthermore, this section aims to describe some of the behaviours depicted by the new elements containing the form *COVID-19* and which occur in non-specialised communication contexts. The migration of terms from specialised to non-specialised contexts points towards a shift in status, from term to non-term. Such change, in some cases, results from determinologisation processes which are analysed further.

Next, in section 6, we analyse the lexical units and terms found in our corpus and describe the respective neological and determinologisation processes.

In section 7, we focus on the lexicographic treatment of four neologisms that have been registered in Portuguese e-dictionaries available in Portugal, namely the *Dicionário Priberam da Língua Portuguesa* (DPLP)[3] and the *Dicionário da Língua Portuguesa* of Porto Editora (DLP).[4]

Finally, based on our corpus analysis workflow, as well as on the systematic comparison of the aforementioned dictionary entries, a template for a lexicographic article targeted at neologisms is put forward in section 8, illustrated by the entry *covid*. This proposal aims, on the one hand, to address the detected inconsistencies in lexicographic representation in the cited Portuguese resources and, on the other hand, to respond to this form's behaviour in the corpus, namely as an element used to create new words (e.g. *covid* + *-ário*).

---

2 https://www.sketchengine.eu/ (last access: 10 June 2022).

3 https://dicionario.priberam.org/ (last access: 10 June 2022).

4 https://www.infopedia.pt/dicionarios/lingua-portuguesa (last access: 10 June 2022).

Overall, the inclusion of neoterms in dictionaries entails several challenges, such as their morphosyntactic classifications, their definition, and which domain label they should or should not be assigned.

2 Neological processes and determinologisation

2.1 Neological processes

In the case of a pandemic, there is an enhancement of neological processes, which emerge more or less spontaneously to quickly resolve communication issues associated with scientific phenomena, which go beyond the understanding of the non-specialised public. New words resulting from these processes are considered to be neologisms. Within non-specialised communication, neologisms may arise from the need to have a communicative impact on the overall community when referring to previously non-existent realities and may even stem from highly specialised contexts. On the other hand, there are neoterms, which designate new specialised concepts produced in a given domain of knowledge. Contrary to a neologism, which is formed spontaneously in relation to communication issues, a neoterm is often formed consciously to designate a concept and distinguish it from others in the concept system to which it belongs, so that it can be used in specialised discourse with a low degree of ambiguity.

Both neologisms and neoterms are linguistic phenomena that are morphosyntactically manifested via the creation of new lexical units, or, semantically, via the attribution of new meanings to already existing lexical units. Neologisms can, therefore, be analysed according to different perspectives. In line with what has been stated by Lino, "neologisms are simultaneously a manifestation of the evolution of a language and the evolution of knowledge, both of which happening at an extremely quick pace" (2019: 10). Terminologists and lexicologists look at the phenomenon of neology from a different standpoint. In terminology science, a neologism is defined as a "term that is specifically coined for a given general concept" (ISO 1087:2019, §3.4.12), whereas in general language, a neologism is defined as "a new word" or "a new meaning of an existing word in the language" (Pruvost and Sablayrolles 2003). The difference is quite significant. That is probably the reason why, in terminology science, the terms "neoterm", "terminological neologism" (ISO 1087:2019) and "neonymy" (Rondeau 1984) have been created to differentiate the conceptual level from the linguistic one. In lexicology and lexical morphology, neologisms are mainly studied as part of word formation and semantics, the latter further exploring topics related to semantic shifting or semantic extension.

Relevant linguistic phenomena include, among others, the formation of terms, the study of collocations and phraseologies, lexical and semantic relations, formal

and semantic neology, as well as variation. In several of these phenomena, the identified linguistic change often has an impact on the dictionary's macro- and microstructure, as well as on the lexical units to be selected to feed the lexicographic resource. These lexicographic activities require that corpora are consistently maintained and up-to-date to detect neologisms and neoterms in time to meet the users' demands.

2.2 Determinologisation processes

Determinologisation (Guilbert 1975, Galisson 1978, Meyer and Mackintosh 2000) is the process by which a term is transformed into a general language word or expression. In these cases, the term does not refer to a concept anymore and, therefore, it is no longer part of a concept system within a given domain. Hence, there is a semantic or conceptual shift prompted by the elimination of one or more essential characteristics of the concept, thereby leading the term to lose its identity and specificity.

Nová (2018: 387) goes further and considers that determinologisation corresponds to the process by which "a scientific term, during its way from a field specialist to a layperson, loses its accuracy, gets new connotations, and the word can be even moved to refer to a completely different thing".

Semantic shift and term variation are the main axes for the study of the specialised lexicon appearing in scientific, technological and technical texts and discourses, both written and oral. Therefore, linguistic change in form and meaning is a dynamic phenomenon that cuts across the entire lexicon. The time-lapse during which meaning is formed, from point A to point B, is recorded in dictionaries, encyclopaedias, vocabularies and ontologies through the choice of lemmas (lexical unit or term) and the definition of the concept and/or the explanation of its meaning.

In terminology, the definition stabilises the relationship between the lexical unit (form) and the specialised concept from a domain of knowledge, in a given period and cultural, political or social context. Meaning is thus fixed in time, which could be short in areas such as science and technology. This finding, resulting from long years of studying the lexicon, allows us to introduce the concept of 'short diachrony', which can be observed in real time. Short diachrony occurs when one observes linguistic change at the level of lexical units, mostly specialised, because of immediate changes in knowledge structures, e.g., when a new concept is introduced in a specialised domain, which has a direct impact on the lexicon. This linguistic change is identified, analysed and classified but it must also be registered, described or defined, and dated in dictionaries to preserve linguistic heritage.

The 'short diachrony' observed in scientific, technical and technological texts contrasts with the 'long diachrony' typically studied in historical linguistics. Short diachrony is of great importance for the construction and update of corpora. In

science and technology, corpora age very quickly when it comes to the study of the lexicon, requiring a constant renewal of the texts that constitute them and constant observance of the published literature in the domains under study. Naturally, there are specialised domains in which more change is observed than in others, with varying rhythms and dimensions.

This is why it is relevant that the Press*CoronaCorpus* corresponds to a monitor corpus, because as stated by Sinclair (1996): "It became clear some years ago that the assumption of a finite limit on a corpus for any length of time was an unnecessary restriction."[5] As depicted in the following sections, it is possible to observe, in a relatively short time span, the appearance and disappearance of lexical units, as well as variation phenomena.

3 Press*CoronaCorpus*: the corpus of analysis

This work has been carried out via the analysis of a dedicated monolingual [EU Portuguese] corpus comprising both a journalistic and a Language for Special Purposes (LSP) subcorpora. As referred to in the introduction, the journalistic corpus – Press-*CoronaCorpus* – was compiled using Sketch Engine, namely the WebBootCaT technology, along with a manual identification of texts publicly available on the internet, to capture newspapers and magazines related to COVID-19 topics. On the other hand, as mentioned above, we also retained official documents produced by healthcare agencies, healthcare professionals and scientists, making up the LSP*CoronaCorpus*. As such, and for the purposes of our study, the journalistic corpus is the corpus of analysis, whereas the LSP-based is the reference corpus. The latter is used to verify semantic shifts, differences in neologism formation, as well as differences in usage, if needed.

The spectrum of Press*CoronaCorpus* is 9 months wide. It is a dynamic corpus intended to represent a snapshot of the language between November 2020 and July 2021 – a time window within the pandemic context – to observe how covid-driven new forms, and corresponding spelling variants, productively entered the Portuguese lexicon.

The texts were gathered during this period on a weekly and sometimes bi-weekly basis, resulting in a large collection of different text types. This collection was organised according to the activities and targeted audience of the texts, culminating in a text typology as systematised in Table 1:

---

5 http://www.ilc.cnr.it/EAGLES96/corpustyp/node19.html #SECTION00090000000000000000 (last access: 10 June 2022).

Table 1: Text typology according to social activities and targeted audience.

Text type	Newspaper	Magazine
Generic	503	
Business & Economics	150	36
Sports	36	2
Environment & Garden		3
Fashion & Socialite		14
Health & Lifestyle		7
IT & Electronics		8
Travel & Culture		24

The number of texts is significantly larger for newspapers when compared with magazines. On the other hand, the most salient type is Generic, whose percentage is 73% when compared with Business & Economics and Sports, whereas the collection of magazines is mostly related to Business & Economics, with 38%.

The capture of media texts on the Internet is not a straightforward task. Due to their increasing subscription-based model, some media pages are not fully available; consequently, the texts were not collected in a balanced quantity throughout the period we have set. To overcome this drawback, we decided to (i) store the collected texts in .txt format, (ii) organise them by trimesters, and (iii) attribute a descriptor to each of them (e.g. PT-NP-GE-2020-11 – which stands for a Portuguese generic newspaper published in November 2020). Such a decision is tied with the manual task of corpus metadata annotation, a process that took place during the compilation of the collected texts with Sketch Engine.

With regard to part-of-speech annotation, we resorted to a tagger embedded in Sketch Engine, specifically the Portuguese FreeLing tagset, since the texts that build up Press*CoronaCorpus* are in EU Portuguese. In short, by merging those two types of corpus annotations, we developed an annotated corpus enriched with metadata, text type and corresponding topic.

Regarding the overall metrics of the corpus of analysis, despite the number of texts not being quantitatively balanced throughout the trimesters, the metrics are considerably robust for the diachronic spectrum under study. As seen in Table 2, the corpus has a little over 40 million tokens and more than 30 million words.[6]

---

6 "A word is a type of token. Words are tokens which begin with a letter of the alphabet" (https://www.sketchengine.eu/guide/glossary/?letter=W) (last access: 10 June 2022).

Table 2: The metrics of the corpus (tokens, words and number of texts).

		PressCoronaCorpus			
		15nov2020-15jan2021	16jan-15mar2021	16mar-15mai2021	16mai-11jul 2021
Tokens	41 610 865	11 163 278	17 052 708	9 557 012	3 837 867
Words	32 930 803	8 849 548	13 438 800	7 640 911	3 001 544
Texts	986	278	405	214	89

4 Exploring the corpus of analysis: the example of COVID-19

For the corpus exploration, we resorted to both simple and advanced queries, depending on the results we were aiming at. Whereas the former use special characters, such as /*/ (e.g. *covid** – from which we obtained matches like *covid*; *covid-19*; *covid19*; and so forth), the latter resorted to Corpus Query Language (CQL), with regular expressions (REGEX) at its core. For instance, to capture *covid* as a monolexical unit without punctuation and digits, we resorted to the following CQL: [word = "*covid*"][!(word = "-.*|.*19")]. The same text mining strategy was used for *Covid* and *COVID*, but with different regexes for the first element [word =], given the different letter cases.

The focus of this paper is the term *COVID-19*, which will be further explored in section 5.

**Table 3:** *COVID-19* spelling variants and frequencies for Portuguese language.

covid-19	10611
Covid-19	5455
covid	3224
COVID-19	1521
Covid	1274
COVID	399
covid 19	37
Covid 19	24
COVID19	11

Table 3 (continued)

covid-19	10611
covid19	8
Covid19	6

Table 3 depicts several spelling variants found in the corpus, with *covid-19* as the most frequent. Interestingly, despite being the form officially used in texts written by experts (much like OMS for EU Portuguese and WHO for English), *COVID-19* has a minor representation (6,7%) in the corpus when compared to *covid-19* (47%), as represented in Graph 1.
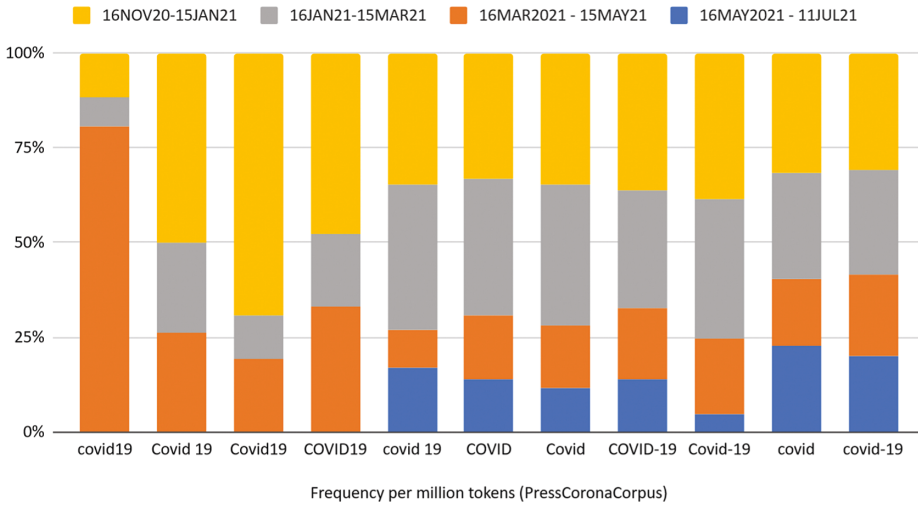
COVID-19 spelling variants

COVID
1,8%
Covid
5,6%
COVID-19
6,7%
covid
14,3%

covid-19
47,0%
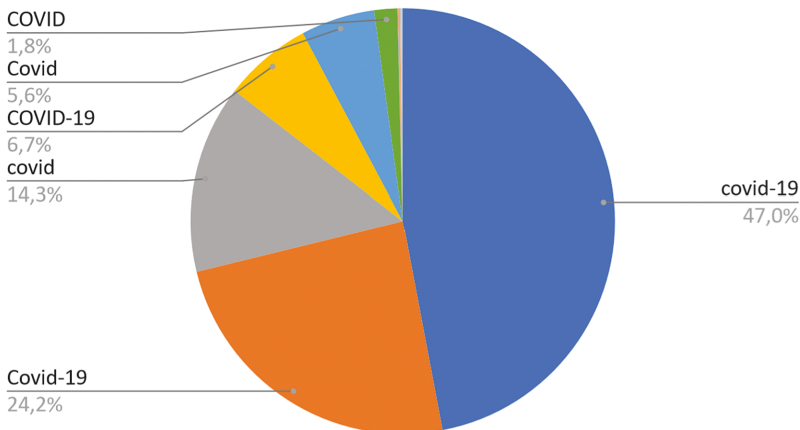
Covid-19
24,2%

Graph 1: *Covid-19* vs *COVID-19*.

The remaining spelling variants, in turn, not only have a low number of occurrences, but also a short period of evidence throughout time. This can be observed in the diachronic spectrum of the corpus, as represented in Graph 2.

Graph 2 depicts the diachronic distribution of the *COVID-19* spelling variants between November 2020 and July 2021. The spelling variants are systematised by trimester and according to per million frequencies, like the example for *covid-19* in Table 4.

Thus, and as observed in the corpus, some spelling variants of *COVID-19* seem to have a shorter lifespan than others. This is the case of variants that have no evidence in the last trimester (16MAY21-11JUL21). These short-life variants, namely *covid19*, *Covid 19*, *Covid19*, and *COVID19* may reflect the instability of these new lexical units in discourse, given its recent lexicalisation.

Distribution of "COVID-19" spelling variants between NOV2020 e JUL2021

Graph 2: Diachronic distribution of *COVID-19* spelling variants (Nov. 2020 – Jul. 2021).

Table 4: Systematisation of the spelling variant *covid-19* according to the per million frequencies by trimester (Nov. 2020 – Jul. 2021).

{covid}*	16NOV20– 15JAN21	16JAN21– 15MAR21	16MAR2021– 15MAY21	16MAY202 – 11JUL21
covid-19	296.24 per million tokens • 0.03%	296.24 per million tokens • 0.03%	207.07 per million tokens • 0.021%	192.03 per million tokens • 0.019%

To further reflect this short diachrony topic, Graph 3 illustrates the time span of 13 forms starting with *covid-* that we have identified in the corpus.
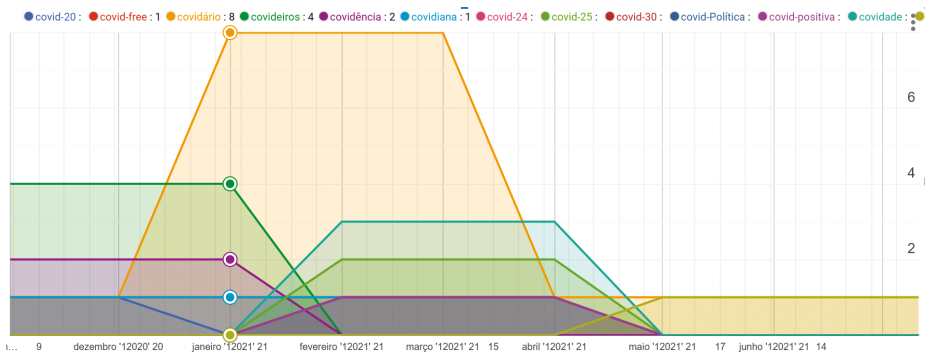
Graph 3: Some examples of lexicalisations with short diachrony/short time span.

Focusing on January 2021 (highlighted with bullets in Graph 3), we can observe that the lexical unit *covidário* has its first occurrence in the corpus in December 2020, and quickly reached its highest frequency in January, with [8] occurrences. This number of occurrences remains until March, finally dropping to [1] occurrence in April and remaining as such until July 2021. On the other hand, the lexical units *covideiros* and *covidência*, with [4] and [2] occurrences respectively, do not appear further beyond February – an evidence over which we hypothesise a short time span, given the diachronic spectrum of the corpus. The lexical units *covidade* [3], *covid-25* [2] and *covid-positiva* [1] are 3 examples with a short time span, which begins in February and ends in April. Finally, the lexical unit *covidiana* [1] is the form denoting the longest time span among the forms under focus here, i.e., it occurs between November 2020 and May 2021. It should be noted that November is not the first attestation of the form, but corresponds to the onset of the corpus compilation.

5 COVID-19: general considerations on word formation

The first word to be analysed is *COVID-19*. *COVID-19* is an acronym formed by the initialisms of the constituents of the polylexical term **co**rona**vi**rus **d**isease 20**19**. The acronym results from a truncation process: *co-* and *vi-*, truncated elements, plus the initial *d* from disease. The number *19* points to 2019, the year when the World Health Organization (WHO) first learned about this new virus on 31 December 2019, following a report of a cluster of cases of an unknown pneumonia disease in Wuhan, People's Republic of China.[7] The WHO's proposal complies with the best practices recommended in 2015 by this organisation for the designation of new human infectious diseases.[8] The English acronym has been imported into Portuguese, despite its complete lack of connection to the corresponding Portuguese polylexical term, which is *doença do coronavírus 2019*. *COVID-19* is, therefore, a hybrid acronym formed by the initials of the constituents of the English polylexical unit and the reduction of the 2019 number, which becomes stabilised in speech and language through a lexicalisation process. The speaker assimilates the acronym as a monolexical unit and integrates it into his/her lexicon functioning as a noun, as attested in our corpus:

---

7 https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/ q-a-detail/coronavirus-disease-covid-19 (last access: 10 June 2022).
8 https://www.who.int/publications/i/item/WHO-HSE-FOS-15.1 (last access: 10 June 2022).

(1) **a** *COVID-19 veio complicar ainda mais a adequada gestão das doenças raras* [COVID-19 **(f.n.)** further complicated the adequate management of rare diseases]

(2) *A presença física no escritório nunca foi obrigatória para a generalidade dos nossos colaboradores ao longo dos últimos oito meses de coexistência com* **a** *COVID-19* [Physical presence at the office was never mandatory for most of our employees over the last eight months of coexisting with COVID-19 **(f.n.)**]

(3) *Será possível exigir testes de diagnóstico para* **a** *COVID-19 nos estabelecimentos de saúde* [The requirement of diagnostic tests for COVID-19 **(f.n.)** will become possible in health facilities]

Given that *COVID-19* refers to a disease (*doença*, in Portuguese), the term's gender in Portuguese is feminine and behaves mostly like a noun in discourse. Nevertheless, the use of the term in the masculine gender also occurs, as documented in the DPLP (cf. Figure 1) ("É também usado como substantivo masculino." [It is also used as a masculine noun.] (DPLP, 2021), since the speakers, by metonymy, designate the disease after the virus causing it, which also happened with other diseases (e.g., the Zika virus and the Zika disease):

COVID-19

(acrónimo do inglês <u>corona</u><u>vi</u><u>rus</u> <u>disease</u> <u>2019</u>, doença de coronavírus 2019 [ano em que a doença foi identificada pela primeira vez])

nome feminino

[Medicina] Doença infecciosa respiratória, causada pelo coronavírus SARS-CoV-2, cujos sintomas podem incluir febre, tosse, dificuldades respiratórias e cansaço, e que, em alguns casos, pode progredir para pneumonia ou falha respiratória.

Nota: Também se escreve com minúsculas (*covid-19*). É também usado como substantivo masculino.

Figure 1: Lexicographic article *COVID-19* in DPLP.

Our research has identified several new lexical units integrating the initial acronym, whereby *covid-* takes on the role of a formative, thus undergoing a re-categorisation. According to Bauer (2003: 330) formative is "a recurrent element of form which correlates with derivational behaviour in some way and yet cannot be identified with a morph".

Figure 2: The evolution of the disease designation until it appears as a formative.

As represented in Figure 2, *COVID-19* is the acronym of the polylexical unit 'coronavirus disease 2019' which in turn, through an ellipse process, loses the hyphen and the reference to the year. When the acronym *COVID* is written in lowercase (*covid*), the lexical unit is perceived as a noun. The term *covid-19*, a feminine noun, is the most frequent in the corpus, with *covid* being the most productive form from a morphological standpoint. Hence, *COVID* is a highly productive lexical unit, behaving as a base form upon which a set of morphological processes act, giving rise to new lexical units while undergoing re-categorisation processes.

The lexical categories of the term *covid* found in Press*CoronaCorpus* are noun (1) and adjective (2):

(1) O futuro, se a **covid** nos permitir, é sempre risonho, nós temos de olhar para a frente sempre com perspetivas de construção [The future, if covid allows us, is always bright, we must always look forward with constructive perspectives].

(2) Ontem, numa corrida só com quatro atletas – por precauções '**covidianas**', quando nos 5000 marcha, durante muitos minutos, andaram todos juntos [Yesterday, in a race with only four athletes – for 'covidian' precautions, when in the 5000 march, for many minutes, they all walked together].

Regarding word formation, we have observed the typical word processes: derivation and composition. In the case of derivation, words are formed through the addition of a prefix or a suffix to the base, which can be a stem, a theme (stem + thematic vowel) or a lexical unit. In this paper, we focus on the acronym *COVID*, which is the base form of the lexical units we are analysing in section 6. The acronym is a lexical unit formed by the word formation process in which an initialism is read syllabically and is a morphological constituent on which word formation operations take place, thereby allowing the creation of formal and semantic derived monolexical neologisms.

Composition, in turn, is a word formation process that operates by concatenating two or more word stems or two or more words. In Portuguese, there are two types of composition: morphological composition, which concatenates stems according to the principles of morphological word formation, and morphosyntactic composition, according to which properties of syntactic structures and properties of morphological structures are combined. The examples we selected illustrate cases of morphosyntactic compounds: *ala covid* and *doente covid*. These compounds are formed by an adjunct structure, that is, they are constituted by two nouns, with similar behaviour to nominal syntactic structures. The right constituent (*covid* in both examples) functions as a nominal modifier, generating a new lexical unit.

Considering the dynamics of the acronym *covid* and since it is highly productive from a morphosyntactic point of view, we use the term to describe formation of lexical units in which *covid* is an endocentric formative.

6 Analysis of morphosyntactic and semantic formation of the neologisms and neoterms

In this section, we observe the occurrences of the base form *covid* to identify behaviours and regularities in word formation, as well as its associations with other elements (prefixes, suffixes), to verify lexical productivity and determine the semantic component of the elements. Based on the data analysis, we will justify whether these words can be considered neoterms or, on the contrary, if having a term in their formation corresponds to a 'false' intuition. As stated by Lombard/Huyghe/Gygax (2021), the neological intuition is an essential feature of neologisms, which can vary according to the individuals and the regularity of lexical creativity processes.

One of the most regular and productive word-formation processes in Portuguese is derivation. Derivation is distinguished from composition in that, contrary to the latter, there is only one autonomous unit of lexical meaning – the base – to which an affix (prefix or suffix) is added to form a new lexical unit. As examples of derived words, we selected the noun *covidário* [covidarium] (whose occurrence in the corpus remained stable over four months) and the adjective *covidiana* [covidian] (one of the first occurrences identified in the corpus). In these examples, two suffixes *-ário* and *-(i)ano* or *-(i)ana* (with a feminine inflexion mark) are added to the base *covid-*. These suffixes occurring after the base determine the lexical category of the newly formed nominal base derivatives (*-ário*, which forms nouns, and *-(i)ano*, which is highly productive in the formation of adjectives), and are also carriers of semantic values. *Covidário* denotes a place (i.e. it has a locative value, being a locative denominal noun according to Rio-Torto et al. (2013) where certain entities remain or are housed, such as in *aviário* (aviary), *berçário* (nursery), *fraldário* (diaper changing room), *infantário* (nursery school) and *solário* (solarium). In turn, the denominal adjective *covidiano* is formed by the adjectival suffix *-(i)ano*, which denotes a living being, as in *bacteriano* (bacterial).

Again, within the processes of derivation, *anticovid* is an example of a word formed by prefixation. The prefix *anti-* combines with the nominal base, *covid*, changing the base word's lexical category (*covid* N → *anticovid* ADJ). Due to its semantic, oppositional value, the prefix *anti-* is combined with bases denoting entities (a disease, in this case), without number inflexion.

As an example of morphosyntactic composition, we chose *covid-drive*, a hybrid compound that occurs twice in our corpus. In terms of its constitution, the noun *covid* is combined with another English-imported noun making up an [N + N] structure. Concerning morphosyntactic compounds, we highlight occurrences such as *ala covid* (covid ward) or *doente covid* (covid patient), which can be included within the so-called "modifying compounds" (Rio-Torto et al. 2013: 93), in that the second element modifies the noun. We are in the presence of an [N + N] structure:

(1) [[ala] N1 + [covid] N2] polylexical unit (compound)

(2) [[doente] N1 + [covid] N2] polylexical unit (compound)

In both examples (1) and (2), the noun *covid* qualifies the N1, allowing us to infer that the *ala covid* denotes a 'place where patients with the disease are housed/hospitalised'. In the second case, *doente covid* denotes a 'person with covid disease'. In these situations, time provides us with the answer of whether these compounds are going to become lexicalised or not and, consequently, if the lexicographer should describe them in a dictionary, which in these cases has already happened. Both *ala covid* and *doente covid* are neologisms, not because the process of word formation is innovative, but because the lexical distribution of their elements has a novelty effect.

From a semantic point of view, these neologisms are formed by lexical units belonging to the lexicon of general/current use, *ala* and *doente*, to which the term *covid* is associated. The question that arises is whether the compound resulting from the combination of these two units forms a neologism or a neoterm. Following our analysis, these units occur in specialised contexts, mostly in public health discourse, but they are not exactly terms because they do not belong to any particular specialised domain. This implies that, in the context of a lexicographic work, these units would not be classified as belonging to medicine, biology, virology, epidemiology or public health through domain labels. In these two examples, the specificity of the < covid > concept is nullified because the characteristics /Respiratory System Disorder/, /Pneumonia/ and /Viral Pneumonia/ are cancelled, thus losing the semantic value associated with specialised domains related to the disease.[9]

A different process is that of the lexical unit *covidário*, which is a derivative formed by analogy, for example, with *berçário* [nursery]. *Covidário* can be defined as an isolated space in a health facility dedicated to covid patients. In this case, the -*ário* suffix does not have a specialised sense and is therefore not a term, being classified as a neologism, although it occurs in specialised contexts, especially those related to hospitals. However, curiously, the DPL, for example, considers *covidário* as a term of Medicine, but does not consider *berçário* as a term belonging to that same domain, despite referring to hospitals in the definition. Have lexicographers been misled by the semantic value of the base *covid*? We find that whenever the formative *covid* appears, there is a tendency to consider the new lexical unit as a term [ex: *covidiota*, *covidivórcio*]

Lastly, the *covidiota* occurrence clearly results from a process of word formation in which the specialised sense of *covid* is lost. From our point of view, this is a process of determinologisation. *Covidiota* is divided into *covid* + *idiota*, being a morphological compound, which is used to refer derisively to a person who does not respect general

---

9 http://covidterm.imicams.ac.cn/#/search?isAdvanced=false&keyword=covid (last access: 10 June 2022).

safety measures, either voluntarily or involuntarily. This unit is a neologism of form and meaning. Formally, it is a portmanteau that corresponds to the blending of two lexical units in which one of the units is truncated: *id* being the end of the *covid* unit and the beginning of the *idiota* unit. We are faced with a haplology, which corresponds to the elimination of one of two consecutive syllables when they are identical or very similar (see Marquilhas 2014: 28). This neologism has no specialised value.

As we can see, it is not always self-evident whether a lexical unit is specialised or not. The fact that the *covid* acronym and its variants are morphologically productive and dynamic (see Table 5) requires an accurate analysis of each of the cases.

Table 5: The productivity of the *covid* acronym.

LEXICAL UNIT	POS	WORD FORMATION	WORD FORMATION
covid	N	Abbreviation of a polylexical unit	ACRONYM
covidário	N	[**covid**]*base* + [ário] *suf.*	DERIVATION
covidiana	ADJ	[**covid**]*base* +[(i)ana] *suf.*	DERIVATION
ala covid	N	[ala]*N* + [**covid**]*N*	COMPOSITION
covidade	N	[**cov**id]*base* + [**id**ade] *suf.*	DERIVATION [HAPLOLOGY]
covidiota	N	[**covid**]*base* + [**id**iota] N	COMPOSITION [PORTMANTEAU]

The determinologisation process is evident in the formation of new general language words, as shown by the examples in Table 6. Reusing the *covid* acronym as a formative element contributes to the process of determinologisation, since the core meaning is used in a superficial manner.

Table 6: Determinologisation process.

covidiota	[+ neologism] [– neoterm]	[composition | portmanteau]
covideiro	[+ neologism] [– neoterm]	[derivation | suffixation]
covidices	[+ neologism] [–neoterm]	[derivation | suffixation]
covidivórcios	[+ neologism] [–neoterm]	[composition | portmanteau]

These lexical units are used in a covid context, but they neither convey specialised features, nor do they belong to a conceptual system of a domain.

7 Lexicographic treatment of neologisms associated with COVID-19

After the extraction and analysis of neologisms from Press*CoronaCorpus*, we now move on to the lexicographic treatment of four neologisms – *covid*; *covidário*; *covidiana*; *anticovid* – previously selected and registered in Portuguese language e-dictionaries, namely DPLP and DLP.

These two lexicographic resources were selected because (i) they are available online, (ii) they are constantly updated with neologisms, both from general language and specialised language, and (iii) they have a very broad list of headwords, each having more than 100,000 entries.

DPLP is a contemporary Portuguese dictionary with about 133,000 lexical entries, whose headword list comprises general language vocabulary as well as terms from various specialised domains. This resource also offers the possibility of browsing entries in European Portuguese spelling, following the 1990 Portuguese Language Orthographic Agreement, and in Brazilian Portuguese, with and without the changes prescribed by this agreement.

DLP is a monolingual Portuguese dictionary that is integrated into the infopedia.pt service,[10] which provides 30 bilingual online dictionaries in several languages (Portuguese, Portuguese Sign Language, English, Spanish, French, German, Italian, Dutch, Chinese, Tetum and Greek). Following European Portuguese spelling, it has two versions: one according to the 1990 Orthographic Agreement of the Portuguese Language, and the other according to the previous standard, that is, the Portuguese-Brazilian Orthographic Agreement of 1945.

The pandemic, as the term itself implies (pan-, Greek *pan*, all), caused the rapid and simultaneous entry of new words – the aforementioned neologisms – in languages around the world. The urgency to publish neologisms of high daily frequency in real time and the need to satisfy the searches of dictionary users often lead to some rash decisions, not allowing lexicographers appropriate and timely reflection on the phenomena and consequent validation of data. Aware of this problem and of certain limitations, we proceed with the comparison of the lexicographic treatment of neologisms associated with the pandemic crisis, intending to answer the following questions:

1) Do DPLP and DLP register the neologisms detected in Press*CoronaCorpus* and which are currently under analysis?
2) Do DPLP and DLP account for these units as neologisms?
3) Do DPLP and DLP consider *covid* as a formative?

The first step was to check whether these units occur in DPLP and DLP. We found that the above-mentioned units are attested in both dictionaries.

---

10 https://www.infopedia.pt/ (last access: 10 June 2022).

As we can see in Figure 3, the neologisms *covidário* and *covidiano* show morphological structures specific to Portuguese, that is, they are considered words derived by suffixation (suffixes *-ário*; *-(i)ano*).

DPLP

co·vi·dá·ri·o
(*COVID[-19]* + *-ário*)
nome masculino
[Medicina] Local, devidamente isolado e equipado que, num estabelecimento de saúde, se destina ao atendimento e ao tratamento de doentes com suspeita ou confirmação de infecção por COVID-19 (ex.: *o hospital criou um covidário para reduzir os riscos de contágio*).

co·vi·di·a·no
(*COVID[-19]* + *-iano*)
adjectivo
Relativo a uma doença infecciosa respiratória causada pelo coronavírus SARS-CoV-2; relativo à COVID-19 (ex.: *confinamento covidiano*). = COVÍDICO

DLP

covidário
co.vi.dá.ri.o   •   kɔvi'darju
nome masculino
MEDICINA área, em hospital ou outra instituição de saúde, especialmente preparada para o atendimento e tratamento de doentes com COVID-19, bem como de indivíduos suspeitos de estarem infetados pelo vírus que origina essa doença

🏛 De *COVID[-19]+-ário*

covidiano
co.vi.di.a.no   •   kɔvi'djɐnu
adjetivo
relativo à COVID-19 (doença respiratória viral causada pelo coronavírus SARS-CoV-2)

🏛 De *COVID[-19]+-iano*

Figure 3: Lexicographic articles *covidário* and *covidiano* (DPLP, DLP).

The first lexicographic data analysed pertains to the formation of words. This information is shown in DPLP in italics between brackets, below the syllabic division of the words. DLP, on the other hand, makes use of an icon to show information about word formation, thereby requiring the user to hover the mouse cursor over the icon to access this information. The two lexicographic resources coincide in the analysis: *COVID-[19] + -ário*. The first point that catches our attention is the fact that this information does not indicate the possibility of *covid* being treated as a formative (*covid-*) for new words, indicating instead that these words – *covidário* and *covidiano* – are formed from the original acronym rather than the *covid* noun itself.

Another topic, often controversial in lexicography, lies in the use of the *Medicina* [Medicine] domain label in the *covidário* article in both dictionaries – especially when confronted with the word *covidiano*, which does not have any domain label. We may question whether this word belongs, in fact, to the medical domain or whether it constitutes a process of determinologisation. Even so, this topic goes beyond our scope, since domain labels in general language dictionaries, in many cases, only function as mere identifiers or word sense disambiguators. Moreover, the dictionaries do not inform us about the criteria for using this label, so we could only make assumptions regarding their use.

Moving on to the illustrative example of derivation by prefixation (starting with the prefix *anti-*), we have the *anticovid* entry. Although our corpus shows the

hyphenated spelling *anti-covid*, according to the Portuguese orthography, the word must be written in an agglutinated form, as evidenced by the entries in the dictionaries. However, while DPLP registers *anticovid*, DLP registers *anticovid-19*, where once again the word formation points to the *COVID-19* acronym (see Figure 4).

DPLP

an·ti·covid

(*anti-* + *COVID*)
adjectivo de dois géneros e de dois números
Que se destina a combater ou proteger contra a COVID-19 (ex.: *máscara anticovid; vacinas anticovid*).

DLP

anticovid-19

ɐ̃tikovid dəzɐˈnɔv(ə)

adjetivo invariável

que visa combater ou prevenir a COVID-19 (doença respiratória viral causada pelo coronavírus SARS-CoV-2)

🏛 De *anti+COVID-19*

Figure 4: Lexicographic article *anticovid* (DPLP, DLP).

We will now turn to the analysis of the units denoting the 'severe acute respiratory syndrome' *COVID-19* (see Figure 5). The orthographic forms are treated differently in each dictionary: the lemma registered in DPLP is the acronym in uppercase (*COVID-19*), while DLP chose as lemma the acronym together with its lowercase form, which can be considered as a spelling variant. The *COVID* entry is also present in DPLP, which further notes the lowercase form: "Também se escreve com minúsculas (covid)" [Also written in lowercase (covid)].

DPLP

COVID

(redução de *COVID-19*)
nome feminino
[Informal] [Medicina] Doença infecciosa respiratória, causada pelo coronavírus SARS-CoV-2, cujos sintomas podem incluir febre, tosse, dificuldades respiratórias e cansaço, e que, em alguns casos, pode progredir para pneumonia ou falha respiratória. = COVID-19
Nota: Também se escreve com minúsculas (*covid*). É também usado como substantivo masculino.

COVID-19

(acrónimo do inglês *coronavirus disease 2019*, doença de coronavírus 2019 [ano em que a doença foi identificada pela primeira vez])
nome feminino
[Medicina] Doença infecciosa respiratória, causada pelo coronavírus SARS-CoV-2, cujos sintomas podem incluir febre, tosse, dificuldades respiratórias e cansaço, e que, em alguns casos, pode progredir para pneumonia ou falha respiratória.
Nota: Também se escreve com minúsculas (*covid-19*). É também usado como substantivo masculino.

DLP

COVID-19, covid-19

kovid dəzɐˈnɔv(ə)

nome feminino

MEDICINA doença respiratória causada por um coronavírus (SARS-CoV-2), apresenta sintomatologia variável, desde casos assintomáticos ou formas de intensidade ligeira (cujos sintomas podem incluir febre, tosse, fadiga ou dores musculares) até situações graves (sobretudo em idosos ou pessoas com problemas de saúde preexistentes), que podem evoluir para cenários de pneumonia, falência de múltiplos órgãos e eventual morte; (inicialmente identificada na China, em 2019, atingiu estatuto de pandemia em 2020)

🏛 Do inglês *Corona virus disease 2019*, «doença de coronavírus 2019», (ano em que foi identificado o primeiro surto da doença)

Figure 5: Lexicographic articles regarding *COVID-19* (DPLP, DLP).

Looking at Figure 5, we see that the unit is classified as a feminine noun in both dictionaries, although DPLP also notes the possibility of it being used as a masculine noun: "É também usado como substantivo masculino." ["It is also used as a masculine noun."]. Although the feminine gender is recommended, our analysis of

the corpus actually attests to the fluctuation in grammatical gender. Instead of resorting to the notes field, which often includes very general information, a better structuring of the data should make the information about gender fluctuation appear in the grammatical information field itself. That is, where *nome feminino* [feminine noun] appears, lexicographers should add *nome masculino* [masculine noun], if the intention is to attest to actual usage in corpora, or simply move the note closer to the gender field, since it is purely grammatical information. It should be noted that in the DPLP case, we have two consecutive notes of different nature: one, focusing on the spelling, indicating that the form is also written in lowercase, and another with a grammatical scope, referring to the word's gender.

We now turn to the matter of the domain labels. While DLP classifies *COVID-19* as belonging to medicine, being preceded by the definition, DPLP shows two usage labels: a diaphasic label, *Informal*, and a diatechnical label, *Medicina*. As lexicographers and terminologists, we may assume the editors' intentions by including these two labels of a different nature, even though they may seem contradictory. However, we question whether an ordinary dictionary user is able to understand these labels. In our opinion, the *Medicina* label specifies the specialised domain, in which *COVID* is a term since it denotes a disease. On the other hand, the *Informal* diaphasic label in the *covid* entry may be justified by its distancing, both semantically and formally, from the original concept of < COVID-19 >, moving from its original context in the medical domain to a less specialised and more general context. However, we may question the use of this diaphasic label, since the fact that a given term becomes popularised does not necessarily mean that it starts being used in informal contexts. In the scale usually established between the informal and formal registers, the use of the reduced form *covid* can be situated in a neutral language register, or can even be used in specialised contexts. In any case, we do not see any advantage in the combination of these two labels, which may even confuse the user or raise further doubts.

Lastly, regarding term formation, DPLP highlights the fact that it is an English acronym.

Concluding our analysis, we are now in a position to answer the above-mentioned research questions:

1) DPLP and DLP record neologisms detected in Press*CoronaCorpus* and selected for analysis. It is important, however, to mention that with regard to the suffixal derivation process, at first, we had chosen as possible candidates for analysis *covidência*, *covidade* and *covidização*, due to their strong neological character. However, at the time of writing this paper, none of these lexical units was recorded in the dictionaries, so we had to exclude them.

2) None of the dictionaries have any identifying marks for neologisms or new entries.

3) None of the dictionaries shows *covid* as an element for forming new words, referring instead to the original acronym (*COVID-19*).

8 Template proposal for a lexicographic article

Following a thorough analysis of the extracted terms selected from our corpus along with the lexicographic treatment observed in the two dictionaries under study, we propose a template for the term *covid*. This proposal should bear in mind the following points:

1) the term *covid* should display its original form, both as an acronym and as a noun in lower case;
2) collocations should be included and highlighted;
3) the term *covid* should be analysed as a formative element of new words.

In Figure 6, the entry referring to the disease is presented. In the 'entry' field, the two forms identified in the corpus are registered. The lowercase form has more occurrences (47% as shown in Graph 1), followed by the acronym form. The 'part of speech' field accounts for the category (n. = noun) and the gender (f. = feminine). Gender fluctuation is not considered here, given that the use of the masculine form is to be avoided and its inclusion may confuse the end-user. This information and other related questions about orthographic rules, for example, could be given by links pointing to other lexical resources, such as spelling manuals or orthographic vocabularies, which can help clarify user questions.

The selected domain label points towards the Medical and Health Sciences domain (Costa et al. 2020). Since the lexicographic definition, starting with 'infectious disease', seems to provide sufficient clarification, the information pertaining to the domain label may be hidden from the user but still is useful to retrieve information for lexicographic purposes. In any case, its insertion is justified, as it allows the lexicographer to better control the terminology and future semantic associations made between this term and other related terms. After the lexicographic definition, which should be as objective and simple as possible, there are usage examples extracted from our corpus. In addition to these usage examples, the observed collocations are registered and also illustrated via real usage contexts.

Should there be any general observations, exemplified in Figure 6, these can be supplied under 'note'.

covid-19, COVID-19 *n.f.*

Domain label Ciências Médicas

doença infeciosa causada por um coronavírus (SARS-CoV-2)

Example: *Devido ao surto de* **covid-19** *vindo da variante sulafricana, a Bélgica, país onde a maioria das corridas irá para a frente, pôs em aberto o cancelamento*

Collocation **vencer a covid-19; contrair COVID-19; tratar a COVID-19**

Example: *à medida que começamos a ver o verdadeiro impacto dos tratamentos adiados em 2020 e os efeitos de longo prazo sobre aqueles que contraíram covid-19".*

Note 1. Acrónimo inglês 'coronavirus disease 2019' , 'doença do coronavírus 2019'.

Figure 6: Lexicographic article regarding *covid-19* and *COVID-19*.

Lastly, since *covid* is used as an element which forms new words, its entry as a formative is also presented in Figure 7. Following the entry, which ends in '-' in this case, clearly showing that this is a word formation element, there is respective grammatical information, as well as the domain label and related sense. The examples are depicted after that. Again, if other types of information are needed, the 'note' field can be used.

covid- *elemento de formação*

Domain label Ciências Médicas

que exprime a noção de covid (doença)

Example: *covidário; covidente; covidiano*

Note 1. Da forma reduzida covid-19, pelo acrónimo inglês coronavirus disease 2019 , 'doença do coronavírus 2019'.

Figure 7: Lexicographic article regarding *covid.*

With these two examples, we believe to have shown that a more rigorous and better segmented structuring process of lexicographic data can bring clarity to the entries and, in turn, to end users. The spelling variants (the lowercase noun and the uppercase acronym) are displayed as lemmas from the start, thereby preventing information of identical scope to be dispersed. The option to resort to the domain label does not necessarily mean that this unit is only used in specialised contexts. Instead, it helps to frame the unit within a previously outlined domain taxonomy.

Although corpus-based examples, and mainly collocations, play a key role in helping the user observe those units in real usage contexts, this seems not to be particularly valued by both DPLP and DLP. Notes, too, will always be helpful in providing other types of information which may be useful to the end user. Ultimately, the innovative contribution of our approach is to introduce an entry for *covid* as a formative element.

9 Concluding notes

In this paper, we analysed new lexical units which have arisen in European Portuguese amidst this pandemic situation. These new lexical units – neoterms and neologisms – emerged for two main reasons. Firstly, the experts needed to designate new concepts which appeared within a specialised context. Secondly, there is also the need to transfer information produced by experts to a non-expert audience. This knowledge transfer is carried out not only by experts, especially from public health settings, but also by journalists and other authors, via discourse production. When this knowledge is transferred to a non-expert audience, information is lost, given that the latter does not have the required knowledge to understand the specialised content of that message. On the other hand, journalists, also a non-expert group, interpret

specialised texts and try to reproduce the information, and therefore information is very likely to get lost, in what concerns both rigour and precision.

While terminology use is present throughout these communication scenarios, neoterms sometimes lose their status and become simple neologisms, thereby leading to determinologisation processes, since there is a context shift which entails a loss of their specialised nature. The extremely fast pace at which new units emerge, either neologisms or terms, has a strong impact on the lexicographer's work. The urgency of publishing neologisms of high daily frequency in real time, as well as the need to meet the research requirements of general language dictionary users, often lead to a certain hastiness, not giving lexicographers the opportunity to conduct a thorough analysis and subsequent validation of their data.

On the other hand, dictionaries can be 'descriptive' or 'prescriptive/normative', establishing the model to follow. Prescriptivism is an approach that attempts to determine the rules of correct usage of a language, while descriptivism is an approach that analyses and describes how the speakers of a language actually use it. Concerning the dictionary as a language model, descriptive guidance has become more usual, a process facilitated by the fact that lexicographers can access increasing amounts of corpora to support their descriptions. We maintain this approach, even though we consider that descriptive dictionaries benefit from a certain normative tone (hence we do not consider the occurrences of *covid* as a masculine noun in our proposal). Users ultimately resort to dictionaries to clarify their doubts and to ensure a correct usage of language.

As stated by Nová (2018: 397), "there is probably no universal way to treat determinologized words, but many of them need a special approach". Some fields can be used for this purpose, as is the case of the notes field, as we have shown. However, it is necessary to take into account that we are dealing with general language dictionaries, i.e., the notes should never be too long.

Aware of the difficulty of registering neologisms in general language dictionaries, especially in the context of an on-going pandemic in which hundreds of specialised units are subject to daily analysis, we found that Portuguese lexicographic resources would greatly benefit from presenting information extracted from analysis corpora and gradually accounting for observed phenomena, such as the determinologisation of terms. In this sense, our paper intends to be a contribution to the advancement of national lexicography.

Bibliography

Bauer, Laurie (2003): *Introducing Linguistic Morphology*. Edinburgh: Edinburgh University Press.

Costa, Rute (2017): Les collocations terminologiques. In: *Provas de agregação, Lexicologia, Lexicografia, Terminologia*. Lisbon: FCSH UNL.

Costa, Rute/Carvalho, Sara/Salgado, Ana/Simões, Alberto/Tasovac, Toma (2020): Ontologie des labels de domaines appliquée aux dictionnaires de langue générale. In: Blanco, Xavier (ed.): *La lexicographie en tant que méthodologie de recherche en linguistique. Langue(s) et Parole – Revue de Philologie Française et Romane* (5).

DLP = *Dicionário da Língua Portuguesa* [online]. Infopédia, Porto Editora, 2003–2021: http://www.infopedia.pt/dicionarios/lingua-portuguesa/ (last access: October 5, 2021).

DPLP = *Dicionário Priberam da Língua Portuguesa* [online]. Priberam, 2008–2021: http://www.priberam.pt/dlpo/ (last access: October 5, 2021).

Galisson, Robert (1978): *Recherches de lexicologie descriptive, la banalisation lexicale: le vocabulaire du football dans la presse sportive, contribution aux recherches sur les langues techniques*. Paris: F. Nathan.

Guilbert, Louis (1975): *La créativité lexicale*. Paris: Librairie Larousse.

ISO 1087: 2019: *Terminology Work and Terminology Science – Vocabulary*. Geneva: ISO.

Lino, Maria Teresa (2019): Neologia e neonímia em língua portuguesa. In: *Linha D'Água* (online), v. 32, n. 3, p. 9–23. Brazil: São Paulo.

Lombard, Alizée/Huyghe, Richard/Gygax, Pascal (2021): Neological intuition in French: A study of formal novelty and lexical regularity as predictors. In: *Lingua*, Volume 254.

Marquilhas, Rita (2014): Fenómenos de mudança na história do português. In: Raposo, Eduardo/Nascimento, Maria Fernanda/Mota, Maria Antónia/Seguro, Luísa/Mendes, Amália (eds.): *Gramática do Português*. Lisboa: Fundação Calouste Gulbenkian, 17–48.

Meyer, Ingrid/Mackintosh, Kristen (2000): When terms move into our everyday lives. An overview of de-terminologization. In: *Terminology*, Volume 6–1, Amsterdam/Philadelphia: John Benjamins, 111–138.

Nová, Jana (2018): Terms Embraced by the General Public: How to Cope with Determinologization in the Dictionary? In: *Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts*. Ljubljana, Slovenia, 17–21 July 2018, 387–398.

Pruvost, Jean/Sablayrolles, Jean-François (2003): *Les néologismes*. Paris: Presses Universitaires de France.

Rio-Torto, Graça/Rodrigues, Alexandra Soares/Pereira, Isabel/Pereira, Rui/Ribeiro, Sílvia (2013): *Gramática Derivacional do Português*. Coimbra: Imprensa da Universidade de Coimbra. ISBN 978-986-26-0640-8.

Rondeau, Guy (1984): *Introduction à la terminologie*. Chicoutimi: Gaëtan Morin éditeur.

Sinclair, John (1996): EAGLES. Preliminary recommendations on Corpus Typology. EAG–TCWG–CTYP/P. http://www.ilc.cnr.it/EAGLES96/corpustyp/corpustyp.html