JCOM

COVID-19 AND SCIENCE COMMUNICATION

COVID-19: a metaphor-based neologism and its translation into Arabic

Amal Haddad Haddad and Silvia Montero-Martínez

Abstract

'Coronavirus Disease 2019 (COVID-19)' is the neologism coined in reference to the pandemic disease currently affecting countries worldwide. The World Health Organization (WHO) was the international entity that coined this neologism in all its official languages, Arabic amongst them. However, in mass media, the most commonly used term is 'coronavirus', which is a meronymic denomination. This corpus-based case study aims at giving new insights into the creation of these neologisms in English and their equivalents in Arabic, and to the adequacy of the meronymic use of the term 'coronavirus' in the English and Arabic mass media.

Keywords Health communication; Science and media; Science writing

DOI https://doi.org/10.22323/2.19050201

Submitted: 16th June 2020
Accepted: 29th July 2020
Published: 30th September 2020

Introduction

Since the breakout in China, in December 2019, of the pandemic Coronavirus Disease 2019 (COVID-19), the World Health Organization (WHO) has been monitoring the development of the phenomenon worldwide. On the outbreak of the disease, the summaries and reports by the WHO used to refer to the disease as 'a pneumonia of unknown cause'. For instance, the summary of December 31 reads: "A pneumonia of unknown cause detected in Wuhan, China was first reported to the WHO Country Office in China on 31 December 2019". Afterwards, on January 10, the WHO issued its first guidebook on the disease, developed with reference to other coronaviruses such as SARS and MERS. Since then, the denomination turned out to be 'Novel Coronavirus' also abbreviated as 'nCov'. This term was used because this virus is a new type of the previously discovered coronavirus, a term denoting the crown-like appearance of the surface projections of the virus. Two days later, a different terminological variant was used by the WHO, '2019-nCoV'. On January 30, the Public Health Emergency of International Concern was declared. Finally, on February 11, 2020, the WHO officially designated the disease with yet another variant "Covid-19", an acronym that stands for 'Coronavirus Disease 2019'. On March 11, 2020, the WHO characterised Covid-19

as pandemic and, eventually, almost all the countries in all the continents had to take drastic measures to fight against it.

This pandemic timeline shows that, in a short time, a variety of terminological phrases or compounds (pneumonia of unknown cause, novel coronavirus), acronyms (nCov), abbreviation or short forms (2019-nCoV) have been used to refer to the same phenomenon before the final neologism was created (COVID-19). However, it is also important to distinguish between the name of the disease and the virus causing the disease. In other words, 'COVID-19' is the pandemic disease, whereas 'severe acute respiratory syndrome coronavirus 2' (SARS-CoV-2) is the virus causing that disease.

The Terminology Coordination Unit of the European Parliament published the COVID-19 event based on Frame-Based Terminology [Faber, 2012; Faber, 2015] to explain, at a linguistic level, the factors which affect the development of the disease (Figure 1). The figure clearly illustrates that the COVID-19 is a type of virus considered as a 'natural agent'; however, the name of the virus as stated by the WHO ('SARS-CoV-2') is listed in the possible complications, which has the potential to lead to misinterpretation. This confusion is due to the fact that 'SARS-CoV-2' is not a possible complication but the name of the virus (an *entity*) itself. Therefore, it should be listed under the category of *natural agent*, as a hyponym (*type-of*) of the term 'virus'.

This paper focuses on these phenomena and places special emphasis on the creation and coining of the neologism Coronavirus Disease 2019 (COVID-19) in English and its transfer into Arabic. This paper questions the use of the term 'coronavirus' in mass media, and whether it fulfils its communicative function across those languages. In general, it has been noticed that scientific research refers appropriately to the name of the disease and to the name of the virus causing it, whereas in popular media the meronymic use of the term 'coronavirus' is widely spread in many languages, i.e. the use of the term 'coronavirus' to refer to the COVID-19.

Literature review

Medicine is one of the specialised domains in which neologisms are constantly generated due to the continuous scientific and technical advances in the field, as well as the appearance of new diseases and unknown phenomena. In this context, neologisms are usually inserted in a language to fill a terminological gap [Schneider, 2018, p. 2], which is usually accompanied by a conceptual gap [Schröder, 2017, p. 248]. In other words, neologisms are used to designate new concepts. Neologisms may also have more than one designation and "the choice of one term or another depends on different communicative and cognitive factors" [León-Araúz, 2015, p. 33]. This means that these terms have both a cognitive and a communicative dimension which determines their use in discourse.

According to Cabré Castellví [1999, p. 92], there are three processes to form terms: compounding, conversion and truncation. Independently of the strategy implemented, coining a neologism must be carried out according to the linguistic code of the language in which they are created: "In general, to create a new designation terminology has available the same resources and mechanisms that the general lexicon has to form new words. Terms conform to the kinds of structure the

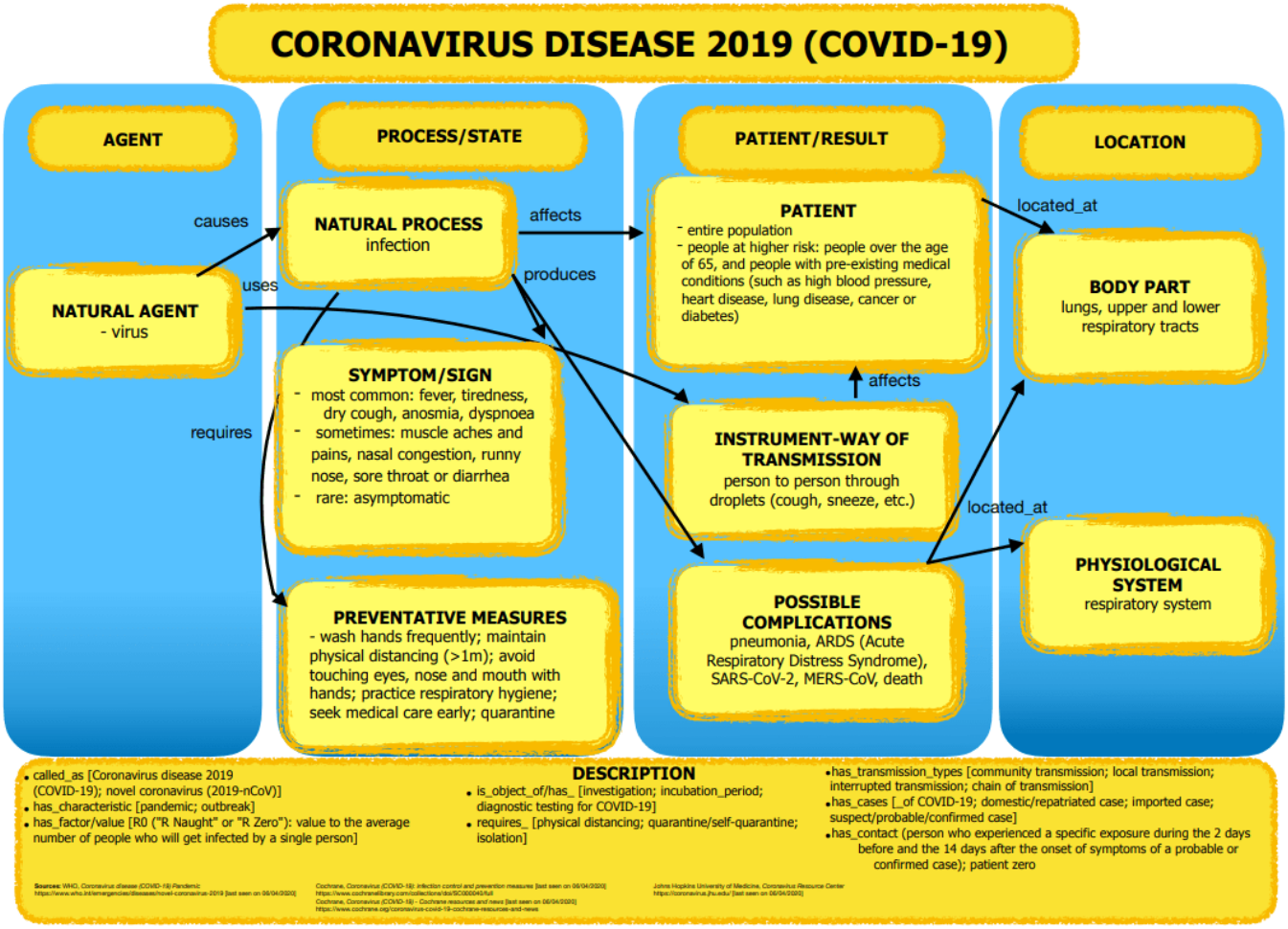

Figure 1. The COVID-19 Event [Terminology Coordination Unit of the European Parliament, 2020].

system allows and are subject to the same rules of combination and restriction" [Cabré Castellví, 1999, p. 92].

One of the pivotal aspects to understand term formation is their form [Cabré Castellví, 1999, pp. 85–86]. Form is related to the constituent morphemes of a term, and varies according to i) the number of morphemes (simple as in *cell*, or complex as in *enzymologist*), ii) the type of morphemes (derived words as in *ulcerous*, or compound words as in *countercyclical*), iii) the combination of words which form a terminological phrase or compound, as in *cardiopulmonary resuscitation*, and iv) the apparently simple units based on complex truncation processes of initialisms, acronyms, abbreviations, and clippings. Truncation is important in term formation and allows to understand this phenomenon in many languages. 'Initialisms' refer to "the combination of the initials of a longer expression" [Cabré Castellví, 1999, p. 86], as in *WHO* which stands for *World Health Organization*. 'Acronyms' are "words formed by combining segments from a developed phrase which are pronounced syllabically" [Cabré Castellví, 1999, pp. 86–87], as in *hazmat* which stands for *hazardous material*. 'Abbreviations' are usually established by consensus and "reproduce a part of a word and practically act as a symbol for the word" [Cabré Castellví, 1999, p. 87], as in *vol* which stands for *volume*. Finally, 'clippings'

or 'short forms' are "based on using the first part of a longer word, or the first word of a phrase" and aim at making discourse shorter [Cabré Castellví, 1999, p. 87], as in *chemo* which stands for *chemotherapy*.

There are different mechanisms to create neologisms, such as meaning extension [Márquez Linares, 2004] and loans [Montero-Martínez, Fuertes-Olivera and García de Quesada, 2001]. In meaning extension, words from the general lexicon acquire new meanings through cognitive processes such as metaphor and metonymy [Márquez Linares, 2004, p. 217], being metaphor "a means of lexical creation both in general and specialized language" [Ureña Gómez-Moreno, 2012, p. 239]. Metaphor-based neologisms are very frequent in specialised languages. They contribute to the understanding and conceptualisation of unknown phenomena, and allow to understand abstract concepts. According to Larson [2011, p. 4], "metaphor is a key element in scientific inquiry because it enables us not only to understand one thing in terms of another but also to think of an abstraction in terms of something more concrete". In the domain of medicine, conceptual metaphors such as *disease is war* are very frequent in discourse construction [Bleakley, 2017; López-Rodríguez and Tercedor-Sánchez, 2017]. Also, metaphor is present in the creation of neologisms [Méndez Cendón, 2004, p. 230], for instance the zoomorphic, metaphor-based term *'dog's ears sign'* and the phytomorphic, metaphor-based *'water lily sign'*. Loans from other languages are grouped under four categories: cultisms, loans from living languages, semantic calques, and pure calques [Montero-Martínez, Faber Bénitez and Buendía Castro, 2011, p. 117]. 'Cultisms' are loans proceeding from Latin and Greek, for instance the term *gland* which has a Latin etymology. 'Loans from living languages' are direct loans from other languages without any modifications, such as the French *bon appétit*, or with a light graphical or phonetic adaptation [Montero-Martínez, Fuertes-Olivera and García de Quesada, 2001, p. 117]. 'Semantic calques' entail meaning extension in a lexical unit, such as in *mouse* referring to the electronic device. In the target language, the use of semantic calques means that "the existing meaning of a word (…) is altered (…) and broadened" [Alberdi Larizgoitia, 2010, p. 18]. Finally, 'pure calques' are foreign lexical units inserted in the target language with formal modifications as the result of a literal translation [Montero-Martínez, Fuertes-Olivera and García de Quesada, 2001, p. 118]. An example is the Arabic term 'العلاج الوظيفي' *'al- ī'lāj al-waẓyfy'*, which results from the literal translation of the English term *occupational therapy*.

Coining neologisms is a challenging task. During this process, terminological gaps are filled in from the conceptual domains of a specific language. Quite quickly, the need to fill the gaps in other languages emerges. Cabré Castellví, Estopà Bagot and Vargas-Sierra [2012, pp. 3–4] highlight that "a language of a whole culture […] needs to have its own terms to be able to express the new concepts of all specialized areas, particularly the most recent ones". Authors like Kajzer-Wietrzny [2011] and Karnedi [2012], from a corpus-based approach, highlight the importance of the role of the translator in shaping national languages. According to Karnedi [2012, p. 1], translators "normally have their own style or language taste which is an important factor in modernizing a language through neologisms". They usually have two options, i) coining neologisms by means of pure calques, ii) use available coining strategies to protect national languages [Kajzer-Wietrzny, 2011, p. 469]. Also, Linder and De Sterck [2016, p. 39], in their study on the reception of English neologisms in scientific discourse by non-native speakers and their equivalents in

other languages, call for the preservation of multilingual science to prevent the impoverishment of national languages within their cultural context. In other words, there seems to be a consensus about the need to prevent the phenomenon of domain loss in the conceptual system of the target languages and cultures when coining terms. For this purpose, Karnedi [2012, p. 1] highlights the importance of language planning, taking into consideration the dominance of English, the *lingua franca*, within the context of global and intercultural communication.

Neologism construction in times of the pandemic

In the case of a pandemic, the need to fill the conceptual and terminological gaps in different cultures and languages is even more urgent, in order to disseminate accurate scientific information and bridge communication between professionals and laypersons. In such cases, the communicative dimension of a neologism is vital. It is necessary to deliver as much information as possible to all individuals, not only to the expert community. Hence, adequate term formation should aim to conceptual comprehension and communicative adequacy.

On the surge of the COVID-19, then referred to as 'a pneumonia of unknown cause' by the WHO [World Health Organization, 2019], the need to fill the linguistic gap and generate a neologism was urgent due to the rapid spreading of the disease. In other words, there was a need to normalise the terminological designation for this new concept with an unambiguos and mono-referential term, that is a term naming just one concept [Cabré Castellví, 1999, p. 108], as well as to avoid the use of terminological variants.

The term proposed and approved by the WHO was COVID-19, which is the acronym of "Coronavirus Disease 2019". According to the WHO report of February 11, 2020, the designation pretends not to make reference to any geographical location, animal, individual or group of people, as well as avoiding being inaccurate or stigmatising: "Guidelines mandated that the name of the disease could not refer to a geographical location, an animal, an individual or group of people. It also needed to relate to the disease and be pronounceable. This choice will help guard against the use of other names that might be inaccurate or stigmatizing" [World Health Organization, 2020].

This neologism is formed by three lexical units: 'coronavirus', 'disease' and '2019'. The meaning of 'disease', according to the *Cambridge Online Dictionary*,[1] is "an illness of people, animals, plants, etc., caused by infection or a failure of health rather than by an accident". The date '2019' refers to the year in which the disease first appeared. Finally, the unit 'coronavirus' is defined in the WHO's website under the 'Q&A on coronaviruses (COVID-19)' section as: "Coronaviruses are a large family of viruses which may cause illness in animals or humans. In humans, several coronaviruses are known to cause respiratory infections ranging from the common cold to more severe diseases such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). The most recently discovered coronavirus causes coronavirus disease COVID-19".[2]

[1] *Cambridge Online Dictionary*: https://dictionary.cambridge.org/.

[2] Available from: http://www.emro.who.int/health-topics/corona-virus/questions-and-answers.html.

According to Peiris [2012], coronaviruses were discovered in the early 1930s when an acute respiratory infection of domesticated chickens was shown to be caused by a virus recently known as avian infectious bronchitis virus. Human coronaviruses were first identified in the mid-1960s. As stated by Peiris [2012, p. 587]: "Until the emergence of SARS in 2003, only two, HCoV 229E and OC43, were recognized as human pathogens. Both were causes of the common cold, considered a mild and insignificant illness and thus not a high priority for intensive research". Consequently, the term 'coronavirus' itself has been in use for quite some time, and cannot be considered a neologism today. This cultism was coined following the etymology of the Latin word 'corona', which means a 'crown' with a curved or circular shape. Mondragón [2020] describes that "under the microscope, these viruses look like round particles surrounded by projections that reminds the solar corona", and that is the reason for their name. This etymological origin reveals the image metaphor the term 'coronavirus' is based on and establishes a resemblance between the real aspect of the virus under a microscope and a crown, or a solar corona. Figures 2 and 3 show the form of avian coronavirus and SARS COV under microscope, respectively. Both figures show the circular shape of the coronavirus and the spikes surrounding it.

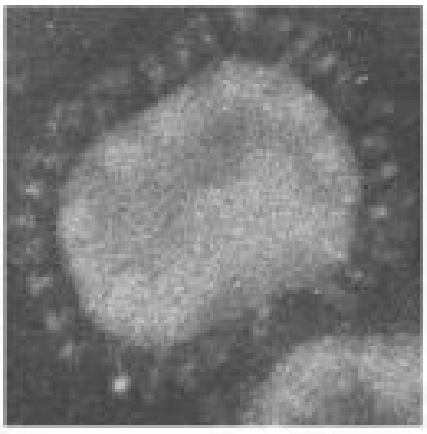

Figure 2. The avian coronavirus (IBV) [Cavanagh, 2005, p. 4].

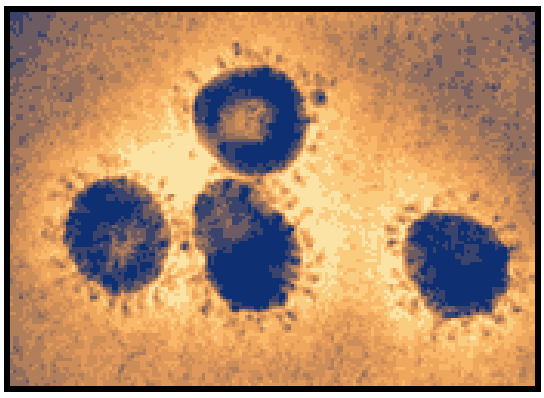

Figure 3. Infectious bronchitis coronavirus virus particles as seen in a colourised electron microscopic image. Image source: F.A. Murphy and S. Whitfield Centers for Disease Control and Prevention (https://www.cdc.gov/ncird/DVD.html).

To appreciate the morphological similarity of Figures 2 and 3 in relation to the solar corona, the basis for the creation of the term 'coronavirus', Figure 4 shows the

visual features around the sun created by magnetic fields. As it happens that the SARS-Cov-2 causing COVID-19 also belongs to coronavirus category, its microscope image in Figure 5 resembles those of the avian coronavirus and the SARS COV. According to Normile [2013, p. 1271]: "all coronaviruses share four "core" genes — the spike, envelope, membrane, and nucleocapsid genes. They also have so-called accessory genes that are scattered through the genome between the core genes". Mondragón [2020] states that coronaviruses "cannot replicate their genetic material themselves and must use cells for this purpose". That is why the virus uses the spikes surrounding it to clip into the cells in order to infect them, spread its genetic material inside and propagate.

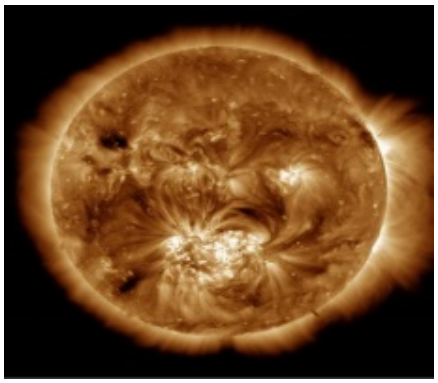

Figure 4. Image of solar corona from NASA's Solar Dynamics Observatory showing features created by magnetic fields. Image credit: NASA (https://spaceplace.nasa.gov/sun-corona/en/).

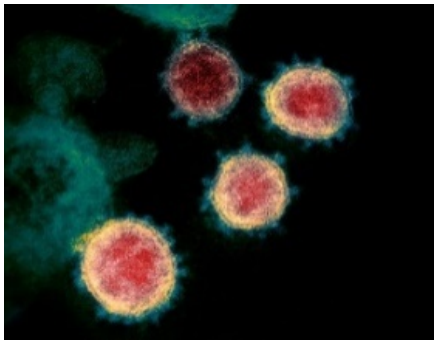

Figure 5. An image of the new coronavirus taken with an electron microscope. Image source: (https://www.nature.com/articles/d41586-020-00660-x).

Specifically, in the case of COVID-19, the SARS-Cov-2 in Figure 5 proceeds in the same way to attack the human cells and expand in the organism. Hoffmann, Kleine-Weber and Pöhlmann [2020] explain how this virus enters the cell: "The spike protein on the virus surface serves as a key for the virus to enter host cells. It facilitates viral attachment to cells and fuses the viral with a cellular membrane, thereby allowing the virus to deliver its genome into the cell, which is essential for viral replication. For this, activation sequences of the spike protein need to be cleaved by cellular enzymes, called proteases".

David Veesler, a structural virologist at the University of Washington in Seattle, states that "understanding transmission of the virus is key to its containment and future prevention" [Mallapaty, 2020]. However, comprehending the scientific facts

and behaviour of the virus is not an easy task, above all for laypeople. For this reason, the metaphor-based cultism 'coronavirus' results helpful to make this information more accessible when talking about any type of coronavirus, as it helps understand and capture the morphological nature of the virus, especially for the non-expert. However, this paper questions the meronymic use of the term 'coronavirus' by mass media to refer to the COVID-19 disease, and the fulfilment of its communicative function across languages, above all in Arabic, a language which does not have a Latin origin and etymology.

Materials and methods

To further analyse and quantify linguistic data concerning COVID-19 in English and Arabic, four different corpora were used and analysed.

In English, an online open corpus of scientific literature about the disease was used. The corpus, obtained from the COVID-19 Open Research Dataset (CORD-19), is publicly available at the Sketch Engine website (http://ske.li/covid_19). It currently has 1,581,577 words (280,762,172 tokens). Also, an English ad hoc corpus of newspaper articles was compiled with Web Search tool, available in Sketch Engine. It has a total of 73,833 words (85,728 tokens). The search criteria included retrieval from prestigious, international electronic newspapers in English, such as *The Guardian (U.K.), The Wall Street Journal (U.S.A.), The New York Times (U.S.A.), The Washington Post (U.S.A.), China Daily (China), The Times of India (India), The Sydney Morning Herald (Australia), The Asahi Shimbun (Japan)*. All of the articles are originally written in English.

In Arabic, two ad hoc corpora were created for this research. The scientific literature corpus was obtained from the WHO's official webpage, as there are no scientific articles published in specialised journals written in Arabic. For this reason, in order to compile a specialised corpus similar to the English one, the World Health Organisation was found the most reliable source. The Arabic scientific corpus has a total of 15,591 words (17,118 tokens). Even though there is a big difference in the total number of words in the English and Arabic scientific corpus, the validity of the linguistic data from the Arabic one is proved in the number of types found in the corpus. In other words, the types refer to the number of different lexical units appearing in a text and does not take into account the repetitions of the same item. The news corpus was compiled with Web Search, restricting the search to international online news websites, such as *Aljazeera* (https://www.aljazeera.net/) (Qatar), *Jordan News Agency* (https://petra.gov.jo/Include/Main.jsp?lang=ar) (Jordan), *Akhbar Alyaowm* (https://akhbarelyom.com/) (Eygpt), *Agence Marocain Presse* (https://www.map.ma/ar/) (Morocco), and *Al Ittihad* (https://www.alittihad.ae/) (U.A.E.). This corpus amounts for a total of 63,603 words (73,919 tokens).

The COVID-19 in English discourse

The keyword search in the English COVID-19 scientific corpus showed the term 'coronavirus' is the second most frequent word in the corpus (Figure 6), with a frequency of 63,685; preceded by 'RNA' with a frequency of 262,087, and followed by 'SARS' with a frequency of 91,971. According to the *Online Collins Dictionary*,[3]

[3] *Online Collins Dictionary*: https://www.collinsdictionary.com/.

the ribonucleic acid (RNA) is "an acid in the chromosomes of the cells of living things which plays an important part in passing information about protein structure between different cells". Conceptually speaking, 'coronavirus' and 'RNA' are linked through a partitive or meronymic relation (*has-part*). In other words, the concept *coronavirus has-part RNA*, the nature of which is of paramount importance in the research studies related to the diagnosis and the creation of vaccines for the disease. This importance is quantitatively reflected in the corpus, where the term 'RNA' is extensively used and appears as the first keyword in the COVID-19 specialised corpus. The third keyword in the list is the term 'SARS', a hyponym (*type-of*) of 'coronavirus'. Clearly, the high frequency of these two keywords shows that the corpus is well compiled and the selection of texts is adequate for the terminological study of the COVID-19 disease.

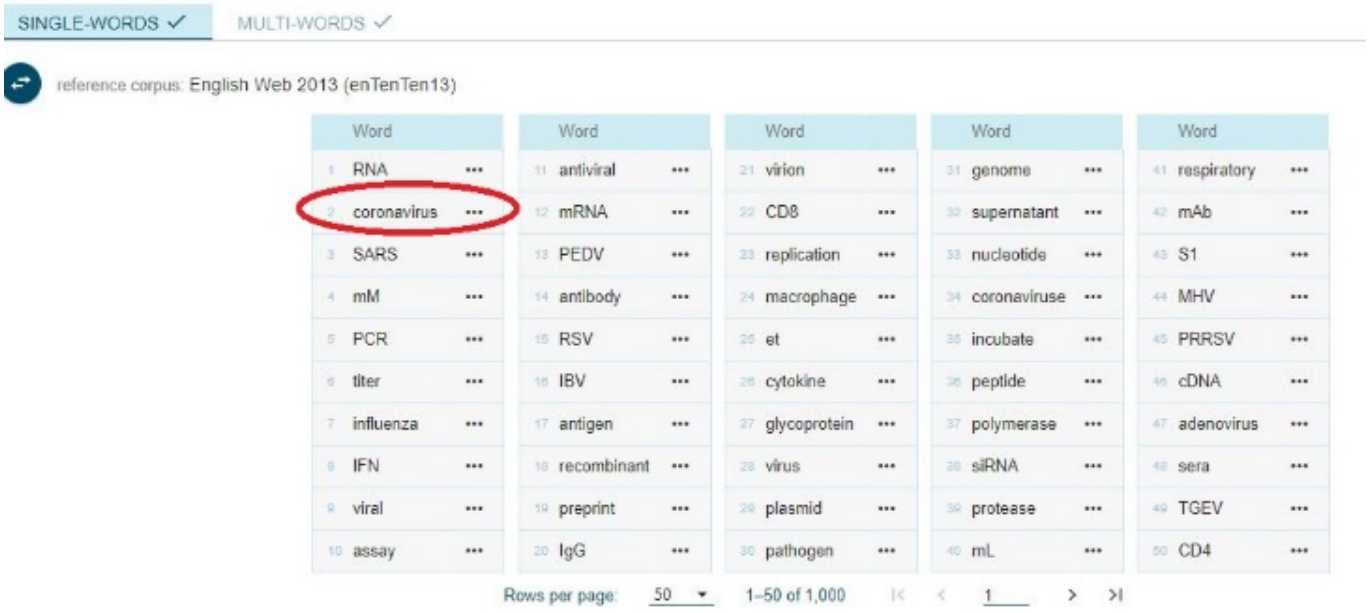

Figure 6. Keywords in the COVID-19 corpus ordered by frequency.

The wordsketches of 'coronavirus' (in Figure 7) show that in almost all cases it co-occurs with other lexical units in the form of terminological noun compounds designating specific types of coronavirus (*type-of*). For instance, modifiers like 'syndrome', 'East' and 'respiratory' form the term 'Middle East respiratory syndrome coronavirus (MERS-Cov)'; 'SARS' + 'coronavirus' ('severe acute respiratory syndrome coronavirus (SARS-CoV)'); 'feline' + 'enteric'+ 'coronavirus' ('FECV'); 'novel' + 'coronavirus' ('nCov'), etc. Also, some compounds point to the *entity affected-by* the *virus* ('bovine coronavirus', 'feline coronavirus').

The concordances list for 'coronavirus' in Figure 8 shows how the modifiers are used in context. Almost all of the results show that this lexical unit refers to an *entity*, a virus which causes (*cause-of*) *disease*, a *process*.

For instance, as seen below in example 1 (a-c), 'coronavirus' is used together with modifiers such as 'novel', 'family' and 'SARS-like'. The resulting terminological compounds point to a *type-of coronavirus*.

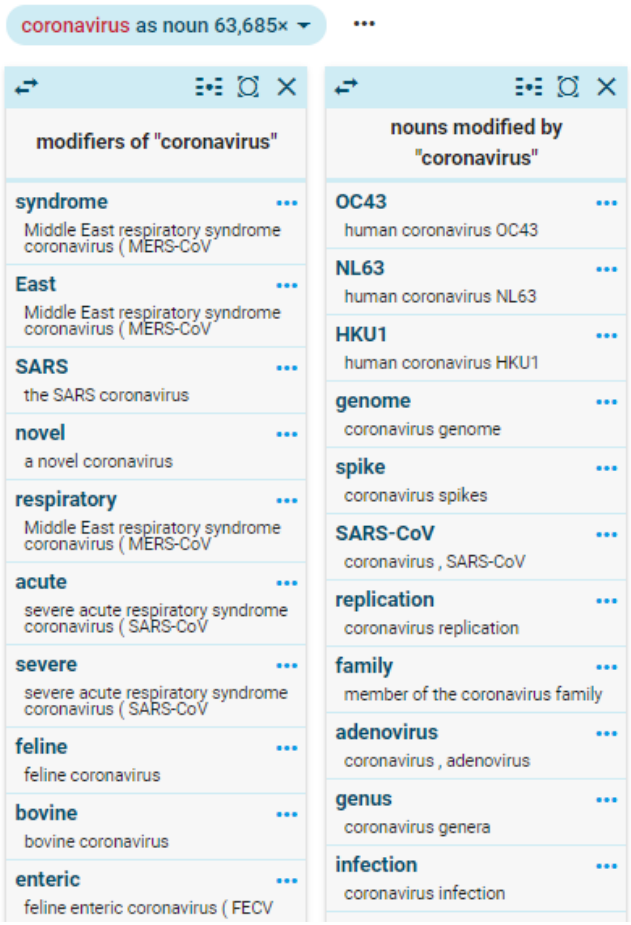

Figure 7. Wordsketches of 'coronavirus' in the COVID-19 corpus.

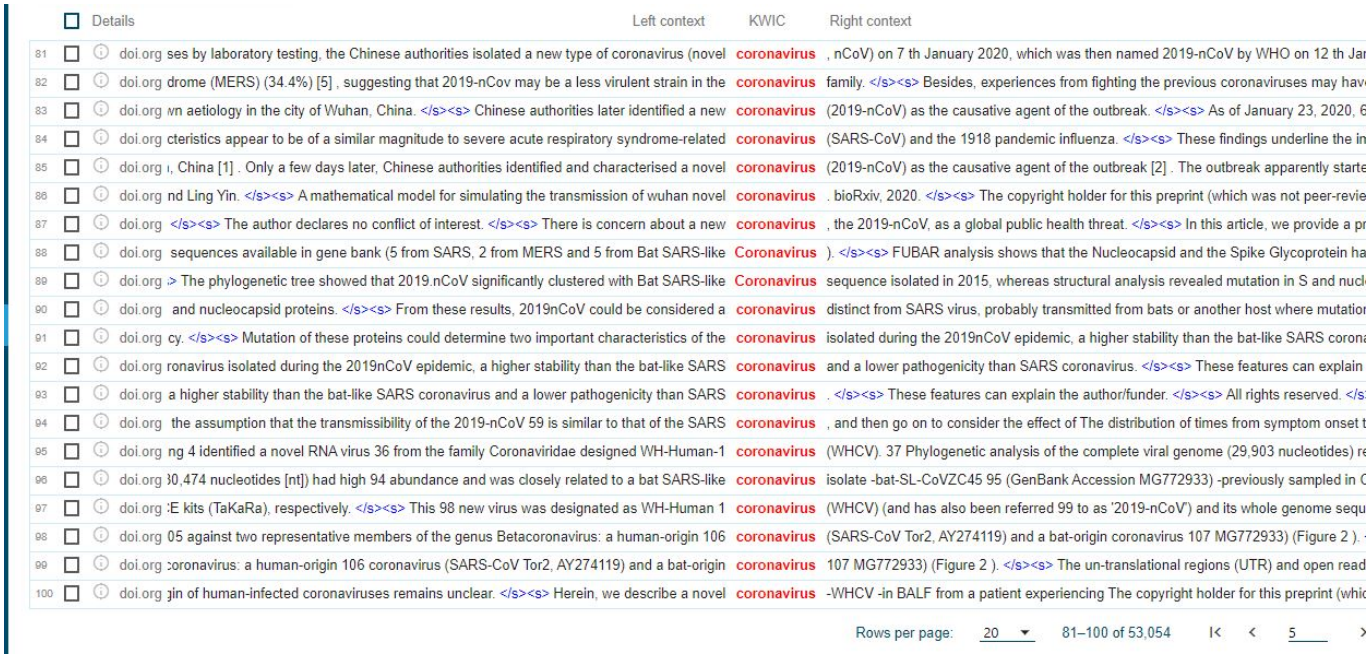

Figure 8. Concordances of 'coronavirus' in the COVID-19 corpus.

Example 1.

(a) "[...] the Chinese authorities isolated a new type of coronavirus (novel coronavirus, nCoV) on 7th January 2020 [...]"

(b) "[...] suggesting that 2019-nCov may be a less virulent strain in the coronavirus family"

(c) "The phylogenetic tree showed that 2019-nCoV significantly clustered with Bat SARS-like Coronavirus sequence isolated in 2015 [...]"

In other words, the concordances found in scientific literature do not refer to the process of the COVID-19 disease *caused-by coronavirus* by using the term 'coronavirus' itself. This terminological use would lack accuracy and would be considered misleading at expert level.

To analyse the situation in mass media, the same procedures have been carried out on the English ad hoc corpus of newspaper articles. In the case of a pandemic, the role of mass media is essential in the distribution of information worldwide. Most countries imposed a lockdown period, in which newspaper readers and online news followers have increased greatly, according to statistics [Hall and Li, 2020]. The keyword search in the press corpus revealed that the lexical unit 'coronavirus', both in small and capital letters, is the most frequent item, with a frequency of 470, followed by the name of the U.S. President ('Trump') appearing 150 times, and the word 'pandemic', with a frequency of 146, as shown in Figure 9.

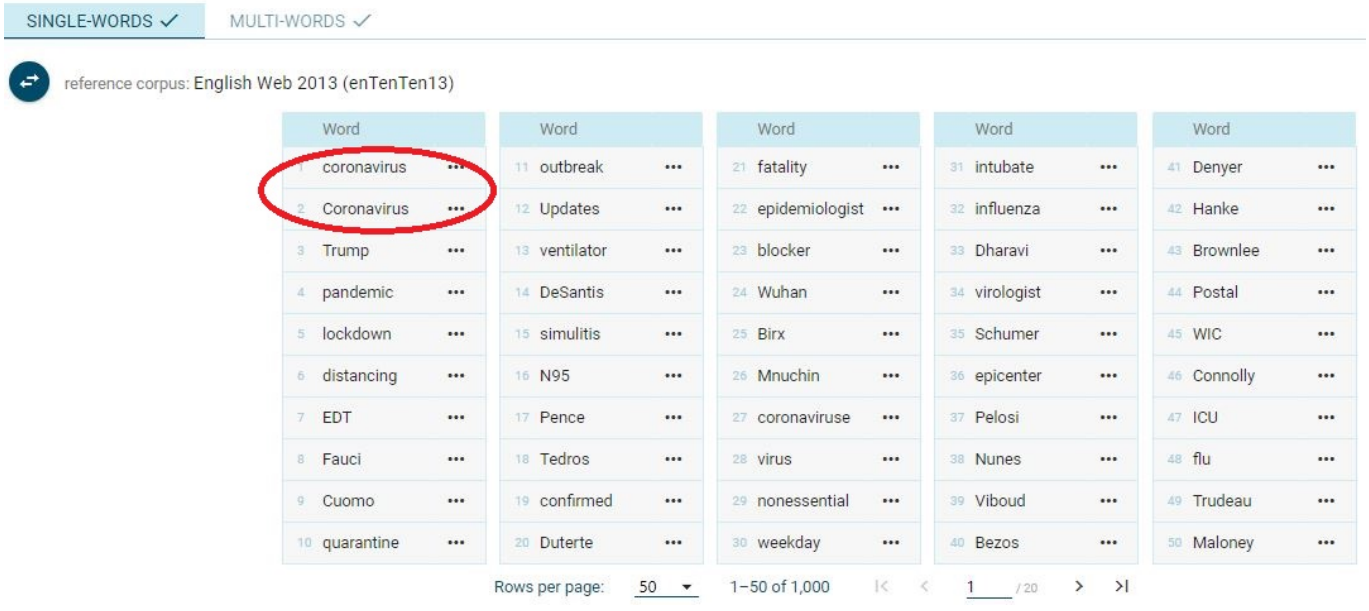

Figure 9. Keywords in the English ad hoc corpus of newspaper articles ordered by frequency.

Concordances of the unit 'coronavirus' show how it is used meronymically in the corpus to refer, in many cases, to the more appropriate term COVID-19 disease (Figure 10).

For instance, in example 2(a), the semantics of the noun phrase "cases of the coronavirus" does not refer to an *entity cause-of disease*, but rather to a *disease*

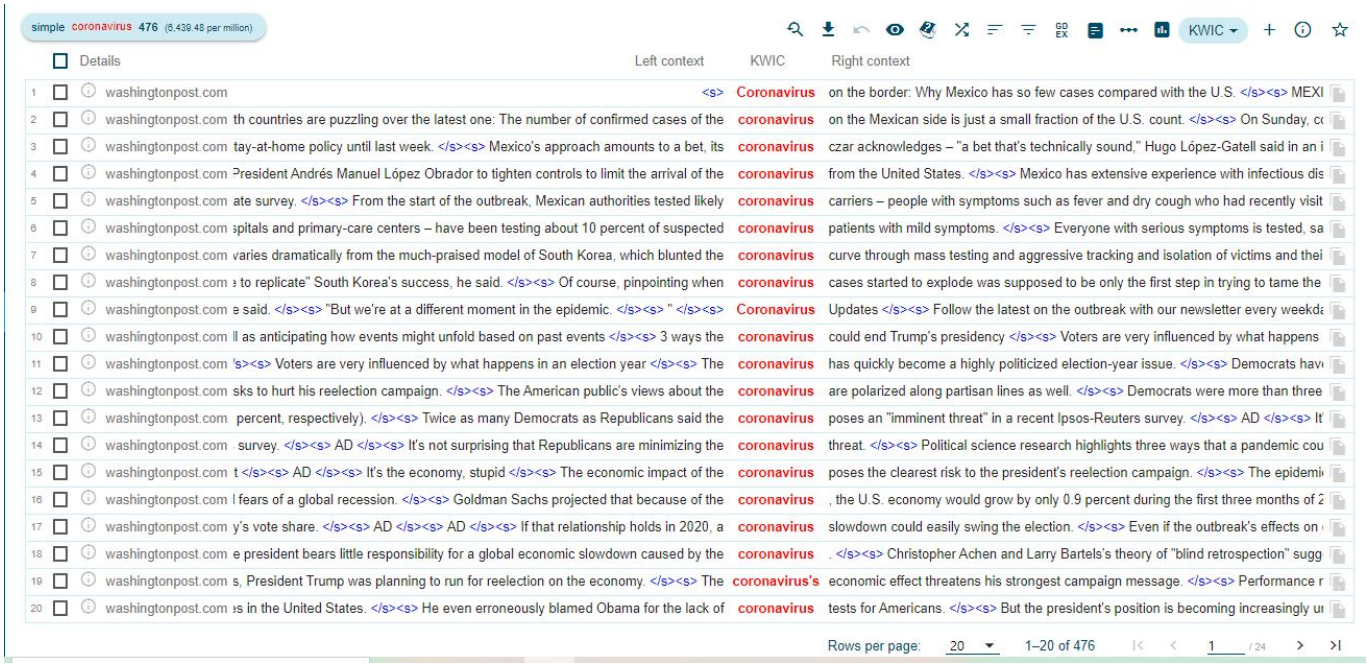

Figure 10. Concordances of 'coronavirus' in the corpus of newspaper articles.

caused-by entity. This can be proved by the inclusion of the conceptual attribute *incidence* ("number of confirmed cases") *attribute-of disease*. It would be impossible to quantify the number of existing virus ("cases of coronavirus"), as separate entities. The logic of the context confirms that quantification here refers to the number of disease processes that have been confirmed. Also in 2(b), "the economic impact of the coronavirus" refers to *disease-incidence affects economy-system*, not to the virus itself as an entity. Finally, the contextual semantics of example 2(c) shows that "coronavirus cases" stands for "COVID-19 cases", a disease process with a beginning and, hopefully, an end. In short, in all these cases 'coronavirus' is used as a synonym for the 'COVID-19' disease.

Example 2.

(a) The number of confirmed cases of the coronavirus on the Mexican side is just a small fraction of the U.S. count.

(b) The economic impact of the coronavirus poses the clearest risk to the president's reelection campaign.

(c) Of course, pinpointing when coronavirus cases started to explode was supposed to be only the first step in trying to tame the pandemic.

The COVID-19 in Arabic discourse

In order to compare the neologism associated to the COVID-19 in Arabic, the ad hoc, specialised corpus containing texts from the WHO's webpage was used. Through a simple keyword analysis in Sketch Engine (Figure 11), it can be observed that the lexical unit 'كورونا' *'kūrūna'* [corona] is the most frequent word in the corpus, appearing 41 times, followed by the negation article 'لم' *'lam'* [no], with

a frequency of 21, the words 'المسبب' *'almusabbib'* [causing] and 'فاشية' *'fāshya'* [pandemic], with a frequency of 20 and 12 consequently, as shown in Figure 11.

Figure 11. Keywords in the ad hoc Arabic scientific corpus ordered by frequency.

The concordances in Figure 12 show that the neologisms coined for COVID-19 in Arabic are 'مرض فيروس كورونا الجديد' *'maraḍ vāyrus kūrūnā al-jadyd'* [new coronavirus], 'مرض فيروس كورونا المستجد' *'maraḍ vāyrus kūrūnā al-mustajid'* [innovated coronavirus], and '٩١ كوفيد' *'kuvyd-19'* [COVID-19]. The three Arabic neologisms were coined by the WHO through a loaning strategy, specifically a pure loan from English, the *lingua franca*, without any phonetic modifications.

The concordances show that the specialised texts by the WHO adequately use the term 'كورونا' 'kūrūnā'*[corona]* to refer to COVID-19 and other diseases produced by a coronavirus, following the WHO's terminological regulation for this concept. For instance, example 3 (a) shows that the unit 'كورونا' *'kūrūnā'* [corona] is used together with 'مرض' *'maraḍ'* [disease], 'فيروس' *'vāyrus'* [virus] and 'الجديد' *'aljadyd'* [new] to refer to the COVID-19. In other words, the conceptual proposition behind this terminological construction is *disease caused-by coronavirus*. Example 3 (b) mentions the term 'كورونا' 'kūrūnā' [corona] in plural in a definitional context in which *coronaviruses* are defined as (*is-a*) a wide range of viruses that may provoke (*cause-of*) illness to (*affects*) animals and humans. Finally, example 3 (c) contains the unit 'كورونا' 'kūrūnā' [corona] (*virus*) as the entity responsible for (*cause-of*) several types of syndromes (*disease*), such as SARS and MERS.

Figure 12. Concordances of 'كورونا' 'kūrūnā' [corona] in the ad hoc Arabic scientific corpus.

Example 3.

(a) [...] مع استمرار تطور فاشية مرض فيروس كورونا الجديد (كوفيد ٩١) تعلن منظمة الصحة العالمية [...]

(ma'a istmrār taṭaur fāshyat maraḍ vāyrus kūrūnā aljadyd (kwvyd-19) tu'lin munaẓamat assiḥa al'ālamya)

[...] With the continuous development of the pandemic new coronavirus disease (COVID-19) the WHO announces [...]

(b) [...] فيروسات كورونا هي فصيلة كبيرة من الفيروسات التي قد تسبب المرض للحيوان وا لإنسان [...]

(vāyrusāt kūrūnā hya faṣyla kabyra min al vāyrusāt allaty qad tusabbib almaraḍ lilḥaywan wal insan)

[...] the coronaviruses are a wide range of viruses that may provoke illness to animals and humans [...]

(c) [...] و من الأمثلة على ذلك ان فيروس كورونا المسبب لمتلازمة الالتهاب الرئوي الحاد الوخيم (سارس) الذي ارتبط بقطط الزباد، و فيروس كورونا المسبب لمتلازمة الشرق الأوسط التنفسية الذي انتقل عن طريق الإبل [...]

(wa min alamthila 'alā dhālik anna vāyrus kūrūnā almusabbib limutalazamat aliltihāb arri'awy alḥād alwakhym (sārs) alladhy irtabaṭa biqiṭaṭ azzabād wa vāyrus kūrūna almusabbib limutalazimat ashsharq alawsaṭ attanafusya alladhy intaqala 'an ṭaryq alibl)

[...] one of the examples is the coronavirus which causes the Severe Acute Respiratory Syndrome (SARS) related to civets and the coronavirus which the Middle East Respiratory Syndrome transmitted through camels [...]

However, despite following the terminological normalisation by the WHO, these transfer processes from English into Arabic of the terms 'coronavirus' and 'COVID-19' result in the loss of the metaphorical function of the cultism

‘coronavirus’. The image association between the virus spikes and the solar corona, the basis for the coining of the neologism, is missed due to the fact that the Arabic language does not have a Latin origin and no resemblance is brought to the mind of the speaker.

The processing of the Arabic news corpus shows that in Arabic media, just as in the English news corpus, the term used to refer to COVID-19 is ‘كورونا’ *‘kūrūna’* [corona], appearing 431 times. Also a misuse of the term COVID-19 in reference to the viral entity (coronavirus) is found. Figure 13 shows concordances for ‘كورونا’ *‘kūrūna’* [corona] in the corpus.

Figure 13. Concordances of ‘كورونا’ ‘kūrūna’ [corona] in the ad hoc Arabic news corpus.

For instance, the context in 4 (a) the meronymic term ‘كورونا’‘kūrūna’ [corona] is used in combination with ‘تفشي’ ‘tafashy’ [expansion] and stands for COVID-19.

Also in 4 (b) and 4 (c) it is equally observed that in ‘أعراض كورونا’ ‘a’rāḍ kānūn’ [symptoms of corona] the sentences does not refer to an ENTITY cause-of DISEASE, but rather to a DISEASE caused-by ENTITY. This can be proved by the inclusion of the conceptual attribute SYMPTOM ‘أعراض’ ‘a’rāḍ’ attribute-of DISEASE.

Example 4.

(a) [...] و بالتالي قد ينقل الفيروس لأشخاص آخرين دون أن يعرف أو يعرفوا إذا خرج من منطقة تفشي كورونا.
(wabittaly qad yantaqil alvāyrus liashkhaṣ dūn an yu'raf aw ya'rifū idhā kharaj min manṭiqa aw min mantiqat tafashy).
[...] consequently the virus may be transmitted to other individuals without knowing whether it came out from the zone of the expansion of corona.

(b) خبر مقلق جديد :أعراض كورونا قد تظهر بعد ١٤ يوما.
(Khabar muqliq jiddan: a'rāḍ kānūn qad taẓhar ba'd 14 yawman).
A worrying news: the symptoms of corona may appear after 14 days.

(c) أعراض كورونا أم نزلة برد .. كيف تعرف؟ و ما الأسلوب البريطاني في الوقاية؟
a'rāḍ kūrūna am nazlat bard ... kayfa ta'rif? Wa ma aluslūb albryṭāny fy alwiqāya?
Symptoms of corona or a cold? How would you know? What is the British method to prevent it?

Results and analysis

After tracing the timeline of the COVID-19 pandemic since its breakout, it has been observed that the WHO referred to the disease with several terminological phrases or compounds ('pneumonia of unknown cause', 'novel coronavirus'), acronyms ('nCov'), and abbreviations or short forms ('2019-nCoV'). Finally, they coined the official name the terminological compound 'Coronavirus Disease 2019' and its acronym 'COVID-19'. This way a terminological gap was filled for a new concept in the virology domain in the English language. The terminological normalisation of such designation by the WHO intended to avoid the use of previous term variants and standardise the use of an unambiguous and mono-referential term, not making reference to any geographical location, animal, individual or group of people, and avoiding inaccuracy and stigmatisation.

The official English term 'Coronavirus Disease 2019' is made up of three lexemes, in reference to i) the type of virus causing the disease, ii) the disease process, iii) the year in which it broke out. The term used to refer to the type of virus causing the disease is the compound lexeme 'coronavirus'. This cultism has been in use for quite some time referring to a large family of viruses discovered in the 1930s. This metaphor-based term alludes to the resemblance found in the virus microscopic morphology, with spikes surrounding it, and the solar corona. Such term contributes to the understanding and conceptualisation of the viral entity and helps to illustrate the virus behaviour and transmission mechanism. This is especially important for laypeople and non-experts whose understanding of abstract concepts may need to be facilitated by accessible constructs based on metaphor. The virus causing the COVID-19 has been named SARS-Cov-2 (Severe Acute Respiratory Syndrome Coronavirus-2).

Examining the COVID-19 English corpus from a terminological point of view, scientific research literature refers to the disease COVID-19 appropriately, and uses the lexical unit 'coronavirus' with modifiers (such as 'Severe Acute Respiratory Syndrome Coronavirus') when designating more specific viral entities; that is,

subtypes of coronaviruses. However, the analysis of the English ad hoc corpus of newspaper articles from international mass media shows 'coronavirus' is not always used in this way. On the contrary, it is widely used to refer to the COVID-19 disease as a meronymic term. This common usage may result in misleading laypeople, as the COVID-19 is not the first disease caused by a coronavirus and is unlikely to be the last. Mass media was a significant source of information about the pandemic. When searching for information on the internet, using the inaccurate term 'coronavirus' in Arabic instead of SARS-Cov-2 or COVID-19, researchers and publics may find search results on avian coronavirus, MERS or SARS-Cov-1. The current pandemic is due to the fact that the SARS-Cov-2 is a human pathogen — that is, it infects humans, a characteristic not shared by most of the coronaviruses until 2003, when the SARS-Cov-1 appeared. To non-expert internet readers, coming across what may seem "contradictory" information may be confusing and hinder the understanding of the virus transmission, a key to containment and future prevention [Mallapaty, 2020].

As the COVID-19 became pandemic in a relatively short time, the English term COVID-19 was used to transfer scientific knowledge in research literature in English but was also transferred to other languages, specially to disseminate information worldwide to the nonexperts. In the case of the Arabic language, corpus analysis shows that the WHO coined the terms 'مرض فيروس كورونا الجديد' *'maraḍ vāyrus kūrūna' al-jadyd'* [new coronavirus], 'مرض فيروس كورونا المستجد' *'maraḍ vāyrus kūrūna' al-mustajid'* [innovated coronavirus], and 'كوفيد ١٩' *'kuvyd-19'* [covid-19]. The strategy to transfer these three neologisms was loaning from English. The first and second are pure calques with formal modifications in Arabic resulting from the literal translation of the original English term 'novel coronavirus'. The acronym is a pure loan, without any phonetic modifications so that it coincides with the English version of the term. Even though the use of calques and loans is very frequent in the scientific and technical fields, in this case they bring the loss of the communicative function of the original metaphor-based image, transmitted through the unit 'coronavirus'. This loan may function appropriately in languages such as Spanish, with a Latin origin where the word *'corona'* is easily understood, but in languages such as Arabic the communicative intention behind the term formation is lost. For many experts, the choice of calques and loans in their native languages accelerates knowledge transfer and makes international communication easier [Montero-Martínez, Fuertes-Olivera and García de Quesada, 2001], the perfect scenario for a pandemic like COVID-19, where the need is even more urgent in order to communicate accurate scientific information and bridge communication between professionals and laypersons. However, the calques and loans also risk the richness and authenticity of the target language, especially if the origin and target cultures do not share a common background of history and believes. In the case of Arabic, the choice most probably works for experts and educated people, but to the common layperson the use of the lexeme 'كورونا' *kūrūna* [corona] will probably fail to facilitate conceptual comprehension of the unknown viral entity responsible for the COVID-19. The concrete reference (the solar corona) for the abstract coronavirus is lost for the Arabic audience.

Also, the Arabic ad hoc corpus obtained from international mass media shows a massive tendency to use the meronymic term 'كورونا' *'kūrūna'* [corona] to refer to the COVID-19 pandemic disease. This extended use is not accurate, as the name of

the virus is easily confused with the name of other diseases caused by other types of coronaviruses. The misuse of the term COVID-19 ‘كوفيد ١٩’ *‘kuvyd-19’* [COVID-19] in reference to the viral entity (coronavirus) was also observed. This lack of preciseness may hinder the correct understanding and flow of information related to the pandemic by laypeople, both of which are essential in the adequate management of COVID-19.

Conclusions

COVID-19 is a neologism which makes reference to the pandemic Coronavirus Disease 2019. The WHO monitored this disease since its breakout and coined the neologism associated to the pandemic. In mass media in English, the use of the meronymic term ‘coronavirus’ is extended worldwide, which may provoke confusion and misunderstanding by laypeople in spite of the urgency of acquiring precise and reliable information. For this reason, using in English the accurate neologism ‘COVID-19’, coined by the WHO, is recommended as the best option to refer to the pandemic in media and disseminate information effectively.

In other languages like Arabic, the neologisms coined are ‘مرض فيروس كورونا الجديد’ *‘maraḍ vāyrus kūrūna’ al-jadyd’* [new coronavirus], ‘مرض فيروس كورونا المستجد’ *‘maraḍ vāyrus kūrūna’ al-mustajid’* [innovated coronavirus] and ‘كوفيد ١٩’ *‘kuvyd-19’* [COVID-19]; nonetheless, the meronymic term ‘كورونا’ *‘kūrūna’* [corona] is very extended, which hinders the correct access to scientific information related to the pandemic. Also, the term ‘كورونا فيروس’ *‘kūrūna vāyrus’* [virus of corona], a loan from English, does not maintain the image metaphor created originally in English. Adequate term formation in Arabic should aim to conceptual comprehension and communicative adequacy, highlighting the need for a culture and language to have their own terms to express new concepts in specialised fields [Cabré Castellví, Estopà Bagot and Vargas-Sierra, 2012, pp. 3–4]. However, neologisms may have more than one designation. The choice depends on the communicative and cognitive factors. Therefore, the recommendation would be to coin an Arabized version of the English neologism, surging from the Arabic conceptual imaginary, to help the native Arabic speakers understand more easily the morphological aspect of the virus and some of its attributes to compensate for the semantic loss produced by the English loans. For instance, a metaphor-based neologism such as ‘الفيروس الشوكي’ *‘al vāyrus alshawky’* [the spinous virus] may be suggested, a term which holds resemblance between the aspect of a coronavirus and a spinous object. This Arabized neologism would co-exist with the English loan to facilitate international communication at expert level, but would also help to maintain the Arabic language and conceptual system updated, without suffering the domain loss phenomenon.

Finally, future research on this issue could take a step further and carry out reception studies among lay readers, speakers with Arabic as their main language and English speakers with a mother tongue derived from Latin. By such studies, their comprehension and use of the COVID-19 neologisms could be analysed to validate or refute the communicative and terminological strategies followed by institutions and mass media.

Acknowledgments

This research was carried out as part of project FFI2017-89127-P, Translation-Oriented Terminology Tools for Environmental Texts (TOTEM), funded by the Spanish Ministry of Economy and Competitiveness. It was also supported by the Investigation Contract from the University of Granada (nº O4430): "Gestión y normalización terminológica, Lanzamiento de UGRCat (másteres y doctorados), Evolución y alimentación de UGRTerm, Creación de recursos TAO para la UGR". Funding was also provided by an FPU grant (FPU18/05327) given by the Spanish Ministry of Education.

References

Alberdi Larizgoitia, X. (2010). 'A typology of calques: the calquing mechanism in contemporary Basque'. *ELUA. Estudios de Lingüística Universidad de Alicante* (24), pp. 13–35. https://doi.org/10.14198/elua2010.24.01. (Visited on 19th April 2020).

Bleakley, A. (2017). Thinking with metaphors in medicine: the state of the art. London, U.K. and New York, NY, U.S.A.: Routledge. https://doi.org/10.4324/9781315389448.

Cabré Castellví, M. T. (1999). Terminology: theory, methods and applications. Amsterdam, The Netherlands and Philadelphia, PA, U.S.A.: John Benjamins Publishing Company. https://doi.org/10.1075/tlrp.1.

Cabré Castellví, M. T., Estopà Bagot, R. and Vargas-Sierra, C. (2012). 'Neology in specialized communication'. *Terminology: International Journal of Theoretical and Applied Issues in Specialized Communication* 18 (1), pp. 1–8. https://doi.org/10.1075/term.18.1.

Cavanagh, D. (2005). 'Coronaviridae: a review of coronaviruses and toroviruses'. In: Coronaviruses with special emphasis on first insights concerning SARS. Ed. by A. Schmidt, O. Weber and M. H. Wolff. Basel, Switzerland: Birkhäuser, pp. 1–54. https://doi.org/10.1007/3-7643-7339-3_1.

Faber, P., ed. (2012). A cognitive linguistics view of terminology and specialized language. Berlin, Germany and Boston, MA, U.S.A.: De Gruyter Mouton. https://doi.org/10.1515/9783110277203.

— (2015). 'Frames as a framework for terminology'. In: Handbook of terminology. Ed. by H. J. Kockaert and F. Steurs. Amsterdam, The Netherlands and Philadelphia, PA, U.S.A.: John Benjamins Publishing Company, pp. 14–33. https://doi.org/10.1075/hot.1.02fra1.

Hall, S. and Li, C. (2nd April 2020). 'COVID-19 proves that media's value is growing — but we need to find better ways to measure it'. *World Economic Forum*.
URL: https://www.weforum.org/agenda/2020/04/covid-19-media-value/ (visited on 19th April 2020).

Hoffmann, M., Kleine-Weber, H. and Pöhlmann, S. (2020). 'A multibasic cleavage site in the spike protein of SARS-CoV-2 is essential for infection of human lung cells'. *Molecular Cell* 78 (4), 779–784.e5. https://doi.org/10.1016/j.molcel.2020.04.022.

Kajzer-Wietrzny, M. (2011). 'Translation of Eurojargon as a source of neologisms in Polish — a corpus based study'. In: Explorations across languages and corpora. Frankfurt, Germany: Peter Lang, pp. 467–480. https://doi.org/10.3726/978-3-653-04563-5.

Karnedi, M. A. (2012). 'The translation of neologisms: challenges for the creation of new terms in indonesian using a corpus-based approach'. *International Journal of Scientific & Engineering Research* 3 (5), pp. 1–13. URL: https://www.ijser.org/paper/The-Translation-of-Neologisms-Challenges-for-the-Creation-of-New-Terms.html.

Larson, B. (2011). Metaphors for environmental sustainability: redefining our relationship with nature. New Haven, CT, U.S.A. and London, U.K.: Yale University Press. URL: https://www.jstor.org/stable/j.ctt5vm557.

León-Araúz, P. (2015). 'Term variation in the psychiatric domain: transparency and multidimensionality'. In: Word formation and transparency in medical english. Ed. by P. ten Hacken and R. Panocová. Newcastle-upon-Tyne, U.K.: Cambridge Scholars Publishing, pp. 33–54.

Linder, D. and De Sterck, G. (2016). 'Non-native scientists, research dissemination and English neologisms: what happens in the early stages of reception and re-production?' *Ibérica* 32, pp. 35–58. URL: https://www.redalyc.org/jatsRepo/2870/287048507003/html/index.html.

López-Rodríguez, C. I. and Tercedor-Sánchez, M. (2017). 'Identification and understanding of medical metaphors by non-experts'. In: Metaphor in communication, science and education. Ed. by F. Ervas, E. Gola and M. G. Rossi. Berlin, Germany: Mouton De Gruyter, pp. 217–246. https://doi.org/10.1515/9783110549928-013.

Mallapaty, S. (2020). 'Why does the coronavirus spread so easily between people?' *Nature* 579 (7798), pp. 183–183. https://doi.org/10.1038/d41586-020-00660-x. (Visited on 1st April 2020).

Márquez Linares, C. F. (2004). 'Polisemia, vaguedad referencial y terminología'. In: Investigar en terminología. Ed. by P. Faber and C. Hurtado Jiménez. Granada, Spain: Comares, pp. 215–226.

Méndez Cendón, B. (2004). 'Estudio descriptivo inglés-español de las metáforas en el lenguaje del radiodiagnóstico médico'. *Panacea* 5 (17-18), pp. 17–18. URL: https://www.tremedica.org/wp-content/uploads/n17-18_tribuna-MendezC.pdf.

Mondragón, T. (23rd March 2020). 'Coronavirus: understanding the enemy'. *Academia Play*. URL: https://academiaplay.es/coronavirus-understanding-the-enemy/ (visited on 1st April 2020).

Montero-Martínez, S., Faber Bénitez, P. and Buendía Castro, M. (2011). Terminología para traductores e intérpretes: una perspectiva integradora. 2nd ed. Granada, Spain: Ediciones Tragacanto.

Montero-Martínez, S., Fuertes-Olivera, P. A. and García de Quesada, M. (2001). 'The translator as 'language planner': syntactic calquing in an english-spanish technical translation of chemical engineering'. *Meta: Journal des Traducteurs/Meta: Translators' Journal* 46 (4), pp. 687–698. https://doi.org/10.7202/003591ar.

Normile, D. (2013). 'Understanding the enemy'. *Science* 339 (6125), pp. 1269–1273. https://doi.org/10.1126/science.339.6125.1269. (Visited on 2nd April 2020).

Peiris, J. S. M. (2012). 'Coronaviruses'. In: Medical microbiology. Ed. by D. Greenwood, R. Slack, M. Barer and W. Irving. 18th ed. London, U.K.: Churchill Livingstone, pp. 587–593. https://doi.org/10.1016/b978-0-7020-4089-4.00072-x.

Schneider, C. (2018). 'Determining survival probabilities for specialised neologisms in medical english and french: a diachronic perspective'. *ASp* (74), pp. 53–76. https://doi.org/10.4000/asp.5350. (Visited on 11th April 2020).

Schröder, M. (2017). 'Broadening the scope of lexical pragmatics: the creation of neologisms in Toposa'. In: Applications of relevance theory: from discourse to morphemes. Ed. by A. Piskorska and E. Wałaszewska. Newcastle-upon-Tyne, U.K.: Cambridge Scholars Publishing, pp. 246–276.

Terminology Coordination Unit of the European Parliament (6th April 2020). 'Image 1: COVID-19 event'. In: COVID-19 event: a frame-based terminology approach. URL: https://termcoord.eu/2020/04/covid-19-event-a-frame-based-terminology-approach/ (visited on 26th July 2020).

Ureña Gómez-Moreno, J. M. (2012). 'Chapter 11. Conceptual types of terminological metaphors in marine biology'. In: Metaphor in use: context, culture and communication. Ed. by F. MacArthur, J. L. Oncins-Martínez, M. Sánchez-García and A. M. Piquer-Píriz. Amsterdam, The Netherlands: John Benjamins Publishing Company, pp. 239–260. https://doi.org/10.1075/hcp.38.18man.

World Health Organization (31st December 2019). *Pneumonia of unknown cause reported to WHO China Office*. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen (visited on 11th April 2020).

— (11th February 2020). *Novel coronavirus disease named COVID-19*. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen.

Authors

Amal Haddad Haddad is a research fellow (FPU) at the Department of Translation and Interpretation of University of Granada (Spain). She is a member in LexiCon research group. Her research focuses on Terminology, Translation and Interpreting, and Corpus Linguistics. She has different publication in national and international journals and publishers. E-mail: amalhaddad@ugr.es.

Silvia Montero-Martínez is a tenured professor at the Department of Translation and Interpretation of University of Granada (Spain). She is a member in LexiCon research group and project coordinator of UGRTERM. Her research focuses on Terminology and Specialized Phraseology, Translation and Interpreting, and Corpus Linguistics. She has published extensively in different national and international journals and publishers. E-mail: smontero@ugr.es.

How to cite

Haddad Haddad, A. and Montero-Martínez, S. (2020). 'COVID-19: a metaphor-based neologism and its translation into Arabic'. *JCOM* 19 (05), A01. https://doi.org/10.22323/2.19050201.

ISSN 1824-2049. Published by SISSA Medialab. jcom.sissa.it