

From *Corona Party* to *Coronapaniek*: Cross-Linguistic Critique of COVID-Related Contributions to Wiktionary

Joachim Grzega

University of Eichstätt-Ingolstadt & VHS Donauwörth

Abstract

Thirteen Wiktionary language versions are analyzed for inclusion of coronavirus vocabulary: How do they include coronavirus-related neologisms? Do coronavirus-related words shape the definition and examples in other entries? If yes, is Wiktionary's neutrality rule obeyed? The results show that neologisms feature particularly in French, English, and Dutch versions. In English, typical examples include proper names, in German quotations from people/organizations and amendments in brackets, and in French war metaphors. Contrary to Wiktionary rules, definitions are often inadequate. Particularly in French, some definitions are discriminatory. Similarly, examples often violate Wiktionary rules and are not always good illustrations of meaning. In Dutch, examples are rarely referenced. Examples reflect ideologies or discrimination particularly in German, French, Dutch, and English. Rule violations are committed even by high-status contributors and occur even with lemmas unrelated to coronavirus, particularly in German and French. In sum, the evidence suggests that collective intelligence does not work (quickly): Wiktionary needs better quality management.

Introduction and Background

Wiktionary, a project of Wikimedia Foundations, is a free-content multilingual general dictionary. Depending on the lemma – or keyword – and the Wiktionary language version, one can get information on the word's meaning, spelling variants, pronunciation, morphological behavior, usage, etymology, synonyms, hyponyms and hyperonyms, example sentences, and translations into other languages, all provided with numerous links. At 18 years past its incipience in December 2002, Wiktionary can be considered "grown up" and well established.

According to lexicographic theory, a lexicographer – that is, a person who compiles dictionaries – is expected to include words entrenched in a language and to include the labeling of a word's pragmatic value (Hausmann, 1989, p. 652). Example sentences should demonstrate the meaning by including many keyword-related words (Harras, 1989; Martin, 1989). Also, while every dictionary is a reflection of its society, lexicographers should understand that in addition to examples with typical evaluative uses, atypical ones should also be included (Harras, 1989).

Anyone can contribute to Wiktionary, but contributors are asked to respect certain rules, some of which are layout-related and some content-related. The following content-related rules are relevant here. Firstly, the Neutrality or "Neutral Point of View" is a general rule of Wiktionary, as well as of Wikipedia: "All encyclopedic content on Wikipedia must be written from a neutral point of view (NPOV), which means representing fairly, proportionately, and, as far as possible, without editorial bias, all the significant views that have been published by reliable sources on a topic. NPOV is a fundamental principle of Wikipedia and of other Wikimedia projects" (URL01). Secondly, definitions (semantic descriptions) are common in all versions. Thirdly, example sentences are encouraged according to the guidelines of many versions, provided the sources are indicated, preferably sources that are academically accepted and

accessible online. Exceptions are the Hungarian, Catalan, Danish and Norwegian versions of Wiktionary, where examples are absent or rare. The German guidelines (as agreed on in 2014) are the strictest and say that, for reasons of space, examples should be used only if a given definition cannot be found in a reference dictionary or Wikipedia. In this case, the *Fünf-Zitate-Regel* ‘five-quotation rule’ should be respected, which says that five quotes from five different authors be given from books or periodicals published in print form and online; the goal is always to illustrate the meaning of the word (URL02, URL03, URL04). The Dutch guidelines encourage example sentences for the illustration of definitions, but, again for reasons of space, only one or two examples; sources are said to be necessary only for sentences longer than 4-5 words (URL05). In the English guidelines, a quotation should illustrate the term’s meaning by the surrounding context, but without being too long: "quotations where that word is the ‘star’ of the sentence serve the reader better than lengthy passages which have the word incidentally buried in them" (URL06). One can also construct example sentences, provided they are not offensive (URL07). The illustration of the word’s meaning is also central according to the French guidelines; yet they explicitly highlight the rule of neutrality, forbid the promotion of ideologies in quotes, and particularly discourage use of quotes from polemic sources and citations that stereotypically discriminate against someone or provoke a rejective attitude in some readers (URL08).

Additionally, the quality of contributions should be maintained by every contributor – at any rate, by those who have been accorded a higher status within the Wiktionary community. On Wiktionary, anyone can contribute, whether they are professional linguists or not, although the Wiktionary administrators (as in Wikipedia) have special rights to ban contributors and block pages. In light of this collective approach, it is interesting to ask how one societally-dominant topic, COVID-19, affects the elaboration of content on Wiktionary.

Research Questions

The impact of a societally "hot" topic (here: COVID-19) on Wiktionary versions are investigated in this study from a cross-linguistic perspective, focusing on (1) the inclusion of lemmas or key words, (2) lemma definitions – in the sense of semantic descriptions, and (3) example sentences.

- RQ1: To what extent do Wiktionary versions include new words related to the coronavirus, but abstain from nonce (single-use) words?
- RQ2: Do coronavirus-related words even occur in the definition of lexemes that are not in any clear semantic relation to the coronavirus?
- RQ3: If yes, are coronavirus-related words defined in line with Wiktionary guidelines and provided with information about their connotations and stylistic use (or jargon)?
- RQ4: Do coronavirus-related words show up in the selection of quotes for lemmas that are not directly semantically related to the coronavirus? If so, for which lemmas?
- RQ5: If coronavirus-related definitions and example sentences are used, do they obey the Wiktionary rules of word-use demonstration, neutrality, and non-discrimination?

These questions are based on the theory that collective intelligence exists, but not as a rapid sustainable phenomenon (cf. Greenstein & Zhu, 2018). Thus it is expected that mistakes, or guideline violations, are not always leveled out immediately, and so will be evident in

coronavirus-related contributions to Wiktionary (up to July 2020, half a year after the coronavirus became relevant for Europe and the Americas).

Literature Review

Features of dictionaries by professional lexicographers are described in part 1 of the HSK handbook on dictionaries (Hausmann et al., 1989). Wiktionary is not (solely) written by professionals. In contrast to its sister project Wikipedia, Wiktionary has been the subject of only a few specific analyses, none of which relate to the research questions of this article. In part 4 of the HSK handbook on dictionaries (Gouws et al., 2013), mentions of Wiktionary are scattered over various articles; they consist in brief descriptions of how Wiktionary works. There has been no analysis on the inclusion and description of neologisms in Wiktionary as a collective dictionary. Neither has there been any research on the impact of societal issues on definitions and example sentences. This is also not a typical issue for research on classical dictionaries.

Cross-cultural studies of Wiktionary do not exist. Neither are Wikipedia versions very commonly the topic of relevant cross-linguistic analysis. Gredel (2018) illustrates convergence across languages as regards certain metaphorical expressions (e.g., the cross-linguistic lemmatization of the coinage "eurosclerosis"). In contrast, Grzega (e.g., 2019a, 2019b) deals with divergences as regards US-EU-Russian relationships and shows that manipulative word use is more prominent in the German Wikipedia than in other language versions. Cultural differences in content are also explored by Callahan and Herring (2011) for articles on famous people in the English and Polish Wikipedias, and by Wagner et al. (2015) on different degrees of including male vs. female personalities in six language versions. Culturally different ways of collaborative authoring are illustrated for French, German, and Dutch by Pfeil, Zaphiris, & Ang (2006). Therefore, cultural differences may also be expected for Wiktionary.

As to ideological bias, we may look, for lack of Wiktionary studies, at recent studies on Wikipedia, which generally exist predominantly for the English, German, and, to a smaller degree, French Wikipedia versions. Since its beginning, Wikipedia has been constantly dealing with criticism concerning information accuracy and neutrality, particularly with respect to persons and political issues. In recent years, Greenstein and Zhu (2018) have claimed that in the long run bias in articles on US politics in the English Wikipedia is in some way neutralized. This means, however, that such bias is not eliminated quickly. Further, slow updating of drug information in 36% of analyzed lemmas in the English Wikipedia was noted by Hwang, Bourgeois, and Seeger (2014). For the French Wikipedia, Bronner (2013) claims that Wikipedia content risks being the result of social consensus, rather than one of truth. For the German Wikipedia, Hoeres (2015) noted that biographies of historians are written in very diverse ways, with their main academic accomplishments hardly playing any role.

Method

The study takes into account the Wiktionary versions in English, Danish, Norwegian, Swedish, German, Dutch, French, Spanish, Catalan, Italian, Russian, Polish, and Hungarian. They are all accessible through the URL pattern <https://xx.wiktionary.org>, with xx representing the language abbreviation (en, da, no, sv, de, nl, fr, ca, es, it, ru, pl, hu), which will also be used in this paper. This selection encompasses English as the dominant European and global lingua

franca, as well as three languages for each cardinal direction of Europe: the north (sv, no, da), west (de, fr, nl), south (es, it, ca), and east (pl, ru, hu). All entries are analyzed in their versions as of 24 July 2020, 23:00 CET. With the help of the search feature of each Wiktionary version, the letter-strings "corona" (for en, fr, de, es, ca, sv, it), "korona" (for pl, hu, da, no), and "корона" and "ковид" (for ru), as well as "COVID" (for all) were searched for in the articles of the Wiktionary versions. The results yielded hits of these letter-strings that cover both lemmas and occurrences in definitions and examples. This includes foreign language lemmas, too; however, for this article, only the non-foreign entries of each Wiktionary version are included. Danish yielded zero hits – it is the only analyzed version whose enlargement had not at all been influenced by the COVID-19 issue as of 24 July, 2020. Despite some quantitative remarks, the analysis is predominantly of a qualitative nature.

Findings for Lemmas

SARS-CoV-2 and COVID-19

All analyzed language versions except for Danish and Norwegian have a lexemic variant for "coronavirus." However, the inclusion was triggered by SARS-CoV-2 only in the Catalan, Spanish, Dutch, Polish, and Hungarian versions; in the other versions, the entry was started much earlier. It is seen as synonymous with the specific coronavirus SARS-CoV-2 only in English, French, and Russian; the others treat it as a hyperonym, or superordinate term, (although, also in these languages, the hyperonym is often used as a synonym for SARS-CoV-2 in general speech). The Russian entry *корона корона*, originally 'crown,' is older, but was extended to SARS-CoV-2 in 2020. "SARS-CoV-2" has entries in Dutch, English, French, German, Polish, and Swedish. The related disease, COVID-19, has entries in English, French, German, Italian, Spanish, Catalan, Swedish, Polish, and Russian. The English, French, and Swedish versions remark that the lexeme for the disease is often used to denote the virus; the English and Russian versions say that the word for the virus is also used to denote the disease.

It may be debated whether it represents a violation of the neutral-point-of-view rule that the entries of the Swedish, French, and English versions connect the virus with Wuhan, China, or both in their definitions of the virus. In the definition of the disease, Wuhan, China, or both are integrated in the Swedish, Italian, Spanish, French, and English versions (but no information on the assumed origin is given in entries for SARS and HIV, which all versions but Spanish have lemmatized). The Italian and Spanish versions also include Korea in their entries for SARS-CoV-2.

In the English Wiktionary version, synonyms for *SARS-CoV-2* are *Chinese virus* and *Wuhan virus* (both labeled "sometimes offensive" and furnished with the comment, "Since the coining of the more formal names *SARS-CoV-2* and *COVID-19* words like *Wuhan flu*, *Wuhan virus*, *China virus* and *Chinese virus* are sometimes considered offensive and tend to be associated with political conservatism in America," as well as an editing remark to the Wiktionary community asking whether this sense should be deleted). Another synonym is *boomer remover* (labeled "slang, humorous") also in the sense of 'COVID-19,' based on the idea that it kills people from the "baby boomer generation."

The Dutch version includes *coronamonster* as a synonym for *SARS-CoV-2*, categorized as pejorative; without special labeling it further includes *wuhanvirus*. The French version

additionally has *SRAS-CoV-2*, labeled Canadian. The other language versions include no synonyms.

Synonyms for *COVID-19* in the English version are: (1) *boomer remover* (labeled "slang, humorous"), (2) *rona* (labeled "informal," clipped from *corona*), (3) *Wuhan flu* (labeled "sometimes offensive," with the same comment as above under *Chinese virus*), and the clipped and rhyming form *Wu flu* (labeled "informal, medicine," which is misleading, since it is typically not a term specific to the professional jargon of medicine), and (4) *Kung Flu* (labeled "humorous, offensive").

As synonyms for COVID-19, the French version has the clipping *cov*, *pneumonie chinoise* (literally 'Chinese pneumonia'), and *pneumonie de Wuhan*, both of the latter without special labeling, although this may have been advisable cross-linguistically due to WHO's 2015 best practices for naming new human infectious diseases without geographical or other group-specific terms (URL09). Another French synonym is *grippette* (literally 'little flu'), labeled ironic.

Coronavirus-Related Innovations with Coronavirus-Related Morphemes

An open online dictionary may include more neologisms than an online dictionary that has a staff of lexicographers behind it who may wait longer to see if a neologism really gets entrenched into a language. At the time of this writing (August 2020), a few activities by professional linguists can already be noted for Catalan (URL10), Dutch (URL11, URL12), English (URL13, URL14), French (Duhamel/Bigorgne 2020), German (URL15, URL16), Italian (URL17), and Spanish (URL18, URL19).

It is noteworthy that the Dutch, French, and English Wiktionary versions are particularly full of coronavirus-related neologisms, while in Italian, Catalan, Polish, and Hungarian they are completely absent. Adjectives for 'against coronavirus' and 'after coronavirus' are ru. *антикоронавирусный antikoronavirusnyj* and *докоронавирусный dokoronavirusnyj*.

Someone infected with the new coronavirus is ru. *коронавирусник koronavirusnik* ('corona virus person'). Someone helping others voluntarily during the coronavirus-related crisis is nl. *coronavrijwilliger* ('corona volunteer') and shows nl. *coronainsats* ('corona commitment'). A hotel that is prepared to accept coronavirus patients is nl. *coronahotel* (alongside nl. *quarantainehotel*). An apéritif via a video conference due to the coronavirus crisis is fr. *coronapéro*. Someone who has to stay in a COVID-affected area is no. *koronafast* ('corona-fast, corona-stuck'), unless it is in Southern Norway to avoid infecting someone in Northern Norway, then the person is in no. *søringkarantene* ('southern quarantine'). Someone responsible for coordinating anti-coronavirus measures in an organisation is listed as *covid manager* in the French Wiktionary, but it terms companies' strategies to use coronavirus-related help operations for marketing effects *coronawashing* (a blend with *greenwashing*). For someone who is skeptical about the sanitary impact of coronavirus, the French version has fr. *coronasceptique*. For someone who claims to know everything about the coronavirus and talks about this in big words the French created *coronologue* (a blend with *philologue* 'philologist' etc.), labeling it as ironic. An attempt to break into a house to steal by telling the resident lies about coronavirus measurements is in Dutch a *coronababbeltruc* ('corona babble trick'). For people who do not respect the official anti-coronavirus rules the German Wikipedia has *Corona-Sünder* ('corona sinner'). In contrast, other versions have a blend of *COVID* and the word for 'idiot': en. *covidiot* (labeled derogatory), ru. *ковидуом kovidiot* (labeled derogatory),

fr. *covidiot* (not further labeled); es. *covidiota* (labeled as Spanish) is lemmatized in the English Wiktionary only; in the German version, the blend is not (yet) included (only suggested for an entry in the corona register [URL20]). In addition, there is nl. *coronakucher* ‘corona-cougher’ in the English Wiktionary (labeled as Dutch). For promenaders inroaching on the compulsory 1.5-meter body distance, the Dutch version has *coronawandelaar* (‘corona promenaders’). For someone who very loudly drives by car through streets that have become quiet due to the coronavirus regulations, there is the Dutch entry *coronaracer*.

For a street that is used to make traffic more fluid during the coronavirus crisis, the lexeme *coronapiste* is given in the French Wikipedia. For ‘fear or hysteria caused by COVID-19’ there is en. *coronaphobia* and nl. *coronapaniek*, as well as en. *coronoia* and nl. *coronahysterie* (the latter two underlining that the fear is unfounded or exaggerated). To express the social, economic, and political turmoil of the coronavirus crisis the English Wiktionary has *coronapocalypse*.

The entries en. *coronavirus party* and *corona party* are given as synonyms of *COVID-19 party* (and its spelling variants), which is defined as a ‘party where people are brought together to spread COVID-19 between each other, so that they may acquire the disease sooner, in order to acquire temporary immunity, as it seems inevitable, and waiting in fear and/or isolation is depressing the economy.’ This definition is supplemented by this statement: "For younger people, the consequence of infection is generally lower, and having immunity releases them from social restrictions that are dampening their lifestyles." Is such a definition not a violation of the NPOV rule?

Coronavirus-Related Innovations without Coronavirus-Related Morphemes

Some innovations related to the coronavirus topic were generated without coronavirus-related morphemes and have already been accepted, predominantly in the French and English Wiktionary (while fully absent in Swedish, Danish, Catalan, and Polish). New cultural greeting habits are denoted by en. *elbow bump*, *elbow shake*, *elbow tap*. The lexeme *social distancing* had already existed as a technical term among sociologists, but entered general language in the wave of anti-coronavirus measures. This is reflected in the form of an Anglicism (fr) or as a loan translation (de. *soziale Distanzierung*, fr. *distanciation sociale*, ru. *социальное дистанцирование* *social'noe distancirovanie*, hu. [without reference to coronavirus, though] *közösségi távolságtartás*) or as a loan rendition (nl. *sociale onthouding*). For the ‘sneezing rule’ (into the crook of the arm) German now has *Niesregel*. If need be, there is hand sanitizer, now also used as an Anglicism in Russian (*санитайзер sanitajzer*). For the 1.5-meter-body-distance rule, the Dutch Wiktionary has *anderhalvemetermaatregel*, for the 14-day quarantine the word *quatorzaine* was given an entry in French. The English *lockdown* (already known in English before COVID-19) is accepted as an Anglicism in the German and Dutch versions; it also occurs in the Russian Wiktionary (*локдаун*), but without specific reference to the coronavirus. In German, there is also the synonym *Shutdown*. The lexemes nl. *quarantainehotel* and no. *søringkarantene* were already mentioned in the preceding section. Similar to fr. *coronapéro* is the martini in quarantine: en. *quarantini*. Someone not wearing a mask is called en. *maskhole* – a blend with *asshole*, classified as derogatory. For someone who believes in what the French microbiologist Didier Raoult suggests as therapies against COVID-19, *raoultiste* and *raoultien* are found in the French Wiktionary. The notion of the heavy distribution of false or misleading news connected to the coronavirus is lexicalized as en. *infodemic*, fr. *infodémie*, it. *infodemia*, and ru. *инфодемия infodemiya*.

For people who are not ill but presumably have the virus, the word *бессимптомник* *bessimptomnik* now exists in general Russian. The corresponding entries de. *asymptomatisch* and fr. *asymptomatique* are linked to sources with *COVID-19* in the title; so is fr. *paucisymptomatique* ‘having few symptoms.’ Before the coronavirus crisis, these words were medical jargon only.

Occasionally, entries are started that seem just individual nonce coinages or just a quote, with or without translation. This is the case with nl. *anderhalvemeterconomie* ‘1.5 meter economy’ (nonce), nl. *gezondheitsoorlog* ‘health war’ (translated quote from French president Macron), nl. *eensaamheidsvirus* ‘loneliness virus’ (quote from Dutch king Willem Alexander), fr. *positive attitude* (from a song title picked up by the daily paper *Le Monde*), and fr. *coronostalgie* ‘nostalgia of the prophylactic lockdown connected to COVID-19’ – both French lexemes are enclosed in quotation marks in the citations given. The French words were added by two contributors who have a superior status within the Wiktionary community ("patroller" and "administrator").

There are no examples of coronavirus-related innovations without coronavirus-related morphemes in the remaining Wiktionary versions.

Synopsis

In brief, we can observe that at least 20 new COVID-related items are included in French, Dutch, and English; German and Russian have 10 (sub-)entries; and in each of the other language versions, there are no more than five new entries. Moreover, even contributors with higher status include nonce words that may never become common in their respective languages, and other contributors do not change these.

Findings for Definitions

Wording and Labeling

In all of the items in the preceding section on lemmas, the definitions include "coronavirus," "SARS-CoV-2," or "COVID-19." But there is more to say. As already pointed out, it may be debated if the lexemes for the virus or the disease should include connections to Wuhan, China, and Korea, especially if information about the region of detection is not included for similar viruses. Such discrimination (which here means use without specific labeling as derogatory) occurs three times in French, and once in Dutch, Swedish, Spanish, and English.

Also, definitions of words should make clear the differences between expert language and everyday language. Wiktionary allows anyone to contribute. It is thus likely that the contributor (and the community, which includes many untrained lexicographers) will miss that there is a difference between a layperson’s and an expert’s use of a word or the semantic change of a simplex word when it is used in a composite word. An example of this is the Dutch entry *coronasterfgeval* ‘coronavirus death case,’ which was started by a user who has contributed actively to coronavirus-related lemmas; the definition given is "(medisch), (neologisme) iemand die is komen te overlijden als gevolg van het besmet geraakt te zijn met een coronavirus" ‘(medical) (neologism) someone who has died as a consequence of an infection with a coronavirus.’ This may be how the word is understood by the majority of Dutch speakers. The definition by the national authorities, however, is "bij de GGD gemelde overleden COVID-19 patiënten" ‘deceased COVID-19 patients registered by the local health

authorities' (URL21). In other words, COVID-19 does not have to be the cause of death, which makes the groups of referents actually larger than in the Wiktionary definition. A good dictionary entry would at least avoid labeling the term "medical" and would ideally point out the discrepancy between how the term is used by national authorities and in which contexts it is used in the media.

Or consider de. *Coronafall* (started by one user) and de. *Corona-Fall* (started by another) 'corona case.' Normally, one graphic variant would redirect to the other, but the two users agree that for the time being they will keep both pages as long as it is not clear which spelling will prevail. The definition given is "Medizin: Auftreten der Erkrankung am Virus SARS-CoV-2 (Coronavirus)" 'medicine: the appearance of falling ill with the SARS-CoV-2 virus (Coronavirus).' The word *Erkrankung* 'illness' is linked to the following definition:

"Störung der körperlichen, kognitiven, sozialen und/oder seelischen Funktionen, die die Leistungsfähigkeit oder das Wohlbefinden eines Lebewesens subjektiv oder intersubjektiv deutlich wahrnehmbar negativ beeinflusst oder eine solche Beeinflussung erwarten lässt"

'disorder of physical, cognitive, social and/or psychic functions which influences the performance capability or the wellbeing of a living being in a subjectively or intersubjectively clearly perceivable negative way or which makes expect such an influence.'

However, *Coronafall/Corona-Fall* in precisely this morphemic construction does not have a definition by the national authorities. Germany's authorities use *COVID-19-Fall* 'COVID-19 case' for a person with at least one positive result of a COVID-19 test (no matter whether it turns out to be a true-positive or a false-positive result) (URL22), and Austrian authorities (*COVID19-)Fall* (URL23) and Swiss authorities use *laborbestätigte Infektion* 'laboratory-confirmed infection,' pointing out that the number of tests is not the same as the number of persons (URL24). There is no specific German word for 'definitely infected person' or for 'person who has fallen ill with COVID-19 symptoms.' Definitions may also differ from country to country, as the WHO warns (URL25). Accordingly, a more appropriate definition of *Coronafall/Corona-Fall* would read 'person who has tested positive for having the coronavirus [for Germany, Austria, and Switzerland; other countries may apply another definition, and definitions may change] [often the word is misinterpreted as definitely infected with SARS-CoV-2 or sick with COVID-19].' Getting one positive test result does not mean that the person is really infected; the quality of the result depends on the test's sensitivity and specificity and on the prevalence of the disease among the group tested (which may actually be very low and thus make the predictive value of positive results lower than 50%). Moreover, getting a positive test result does not mean that the person has fallen sick with clear, typical symptoms. Of course, it may well be that this is confused among laypeople, but a good lexicographic entry would point this out. Actually, it is not clear either which definition is used in the quotes used under the lemmas *Coronafall/Corona-Fall*, either.

One might think that in spring 2020, many Wiktionary versions would have included a word for "coronavirus" or "SARS-CoV-2" or "COVID-19" in the entry for "pandemic." In the English Wiktionary a *pandemic* is "a disease that affects a wide geographical area and a large proportion of the population." The definition is similar in the German, Dutch, French, Spanish, Italian, Russian, Polish, Danish, and Swedish Wiktionary versions (in Norwegian there is no equivalent). It is equally similar for hu. *pandémia*, while the synonym *világjárvány* lacks a definition, thus violating a Wiktionary rule. No version refers explicitly to the WHO as a

defining institution, not even one of those labeling the entry ‘medical.’ Moreover, none has or had in earlier article versions the original use by the WHO, according to which a pandemic also includes many deaths world-wide. The original passage was deleted by the WHO in April 2009 (Doshi, 2011). An ideal entry would differentiate between everyday language and official language and also note the semantic change, especially if the public discusses referents of a lexeme in its old and new meanings. At any rate, the definitions do not include references to coronavirus or COVID-19. However, the French and Dutch entries include quotations referring to coronavirus or COVID-19.

Synopsis

In sum, discriminating definitions occur several times in French entries (even with the incorrect jargon label "medicine"). They appear particularly in the sense that derogatory terms are not labeled as such, giving users a false impression of the connotation of the word. In English, Swedish, Dutch, Italian, and Spanish discrimination is rare; in the other versions it is absent. Furthermore, the label "medicine" is sometimes misused. It seems as if users confuse the jargon/register labels with indicators of the conceptual field. Finally, misleading semantic information (common language vs. medical jargon, old meaning vs. new meaning) is left uncorrected.

Findings for Example Sentences

Preliminaries

In the various Wiktionary versions there were a number of lemmas (including foreign language lemmas) that contained "coronavirus" (or an existing short form) in the example sentences or the title of the sources of the example sentences. If we exclude all foreign-language lemmas and just concentrate on the lemmas in the respective language of each version, then our corpus includes 11 entries in Russian, 8 in Polish, 2 in Italian, 1 in Swedish, 1 in Spanish, and none in Hungarian, Catalan, or Norwegian; more than 50 hits are only available for English, French, German, and Dutch. If one looks for "COVID" hits, the picture does not change significantly. In the latter four Wiktionary versions, example sentences and their references are not equally prominent. In the English Wiktionary one specifically has to click on an icon next to a definition; the click will reveal quote and directly succeeding source; in the French version quote and directly succeeding source are immediately visible when one opens the entry; while in the German and Dutch versions the source is detached from the quote and listed at the bottom of the entry. This section will show that (a) examples do not always demonstrate definitions, (b) examples seem to promote propaganda and ideologies (not only with words somehow related to the pandemic issue, but also with everyday words), and (c) example types show cultural preferences. (In the following sections, the quotes maintain the boldface and italics of the original; translations adhere to the literal phrasing of the original.)

Demonstrating Definitions?

Predominantly, quotations should be selected in order to demonstrate the meaning of a word. Sometimes this is done successfully, sometimes not. In the English Wiktionary, the selected quotation for *case fatality rate* helps –

(1) That 3.4 percent mortality rate, though still fluid, is formidable, even compared next to the **case fatality rates** (the number of deaths compared to the total number of those diagnosed) for other recent outbreaks.

– , while the quote for the synonymic acronym *CFR* is useless:

(2) The **CFR** for the SARS epidemic was around 9 to 10 percent, while MERS, first identified in 2012, is around 35 percent.

Similarly, the demonstration is good for fr. *distanciación sociale*, but not for fr. *social distancing*. One contributor-constructed Italian sentence –

(3) la positività al COVID-19 si rileva con un tampone che misura il virus circolante in gola o nelle narici

‘positivity for COVID-19 is determined with a pad that measures the virus circulating in the throat or in the nostrils’ –

is used in the illustrations of six words: *rileva* ‘determines,’ *misura* ‘measures,’ *circolante* ‘circulating,’ *gola* ‘throat,’ *narici* ‘nostrils,’ and *tampone* ‘pad, pledget.’ Assuming that with each keyword all the other words are understood, it seems that the sentence helps explain *rileva*, barely helps with *tampone*, and for the other four words the sentence’s explanatory power seems moderate.

Promoting Propaganda?

If it happens more than just sporadically that example sentences do not explain words well, they may do something else. Particularly for the German version, the high number of example sentences including a coronavirus-morpheme is surprising, since – as was mentioned before – the German guidelines say that quotations should only be used if the word cannot be linked to a common reference source, including Wikipedia. A Wiktionary help page even offers a list of reference works (URL26). Thus, if a lemma includes a quote although it is established in the Wikipedia or a reference book, there is some suspicion that a quote may be (mis)used for something else than encyclopedic information. It may purvey propaganda or a certain ideology, which would violate the neutrality rule. In the following sections, we will see, firstly, that users do indeed present specific policies in a certain ideological light (with lemmas linked and lemmas not linked to the coronavirus issue), and secondly, that users discriminate against certain communities and even individual people (with lemmas linked and not linked to the coronavirus issue).

It is understandable that a new word cannot be found in printed reference books and that it may be controversially developed as a Wikipedia article. Wiktionary offers a good chance to represent the word in five different quotes according to the German Wiktionary rules. Thus one user (we will call her User-de1) enlarged on the German entry *Lockdown*, an Anglicism, with a coronavirus-related definition and five quotes on 19 March 2020. However, only one quote serves to illustrate the meaning of *Lockdown*, due to its proximity to the synonym *Ausgangssperren* ‘bans on going out.’

(4) "Ich bin kein Freund des *Lockdowns*. Wer so etwas verhängt, muss auch sagen, wann und wie er es wieder aufhebt," sagte Montgomery. Die SPD-Europapolitikerin Katarina Barley sagte, Ausgangssperren sollten das letzte Mittel sein.

“I’m not a friend of the lockdown. Who orders such a thing also has to say when and how to end it,” said Montgomery. SPD EU politician Katarina Barley said bans on going out should be the means of last resort.’

One quote associates the word with *Vollbremsung* ‘full braking,’ but this is an individual metaphorical use by the person interviewed.

(5) "Deutschland braucht eine Vollbremsung, einen *Lockdown*, mindestens so, wie ihn Italien jetzt hat," forderte Stephan Ortner [...].

"Germany needs a full braking, a lockdown, at least the way Italy has it now," Stephan Ortner claimed.’

What this quote and the remaining three quotes do, rather, is to present *Lockdown* as worthy of support. This is also done through a supplementary sixth quote by another user (we will call him User-de2):

(6) Andere fragwürdige Studien und Interviews mit Scheinexperten behaupten, dass es keinen *Lockdown* brauchte oder dass das neue, unbekannte Virus weit weniger gefährlich sei als behauptet.

‘Other dubious studies and interview with pseudo-experts claim that a lockdown is not needed and that the new, unknown virus is much less dangerous than assumed.’

Normally, if not done by the Wiktionary community, at least an administrator should intervene and say that this is a violation of the rules that the community has given themselves. The problem is that User-de1 is an administrator herself. She does not call User-de2’s attention to the rule violation and violates the rule herself. Other administrators frequently viewing and/or contributing to coronavirus-related topics do not correct this either.

Similarly, User-de1 starts a new entry, de. *PCR-Test*, again with five quotes, three of which help to understand the meaning of the word. One quote is such a neutral sentence that it does not illustrate the meaning at all. The remaining quote is this:

(7) In mehreren Youtube-Videos wird behauptet, dass *PCR-Tests* zum Nachweis von Corona-Infektionen in 30 bis 50 Prozent der Fälle falsche Ergebnisse lieferten. Das Ausmaß der Pandemie werde deshalb überschätzt. Der Virologe Christian Drosten hat den Aussagen bereits öffentlich widersprochen.

‘In several YouTube videos they claim that PCR tests for evidence of corona infections yield false results in 30 to 50 percent of the cases. The extent of the pandemic is therefore being overrated. The virologist Christian Drosten has already publicly contradicted these statements.’

The word appears in the first sentence. That sentence would be enough; it would also show that the test is not perfect. User-de1, however, adds two more sentences, by which she changes the negative connotation (maybe also denotation) into a positive one. As a matter of fact, as mentioned above, such a test, depending on the group to which it is applied, could indeed produce these many incorrect results (or fewer, or more). Here, the link to the Wikipedia entry might be enough, or a quote from an overview article from the German medical magazine *Deutsches Ärzteblatt* (URL27) (cf. also Reiss/Bhakdi 2020). With the quote, User-de1 not only gives lexicographic information, but shapes a certain view.

It is not certain that User-de1 did this on purpose, as she also creates entries that include quotations critical of the mainstream opinion. On 18 May 2020, she started the entry *Immunaktivität* and included this quote:

(8) So gesehen, können die derzeitigen Quarantäne-Maßnahmen auch kontraproduktiv sein, etwa, wenn Menschen dabei depressiv werden oder eine Angsterkrankung entwickeln und die antivirale *Immunaktivität* unterdrückt wird [...].

‘Viewed this way the current quarantine measures may also be counter-productive, for instance when people develop a depression or an anxiety disorder and the antiviral immunity activity gets suppressed.’

According to the German guidelines, quotes are not necessary with composite words if both elements do not display an aberrant use. Consequently, also for de. *Hirnschaden*, compound of de. *Hirn* ‘brain’ plus de. *Schaden* ‘damage,’ no citation is required. Nonetheless, User-de2, when beginning the article, includes:

(9) Britische Neurologen haben jetzt schockierende Einzelheiten in der Zeitschrift "Brain" veröffentlicht, wonach SARS-CoV-2 selbst bei Patienten mit leichten Symptomen oder bei bereits Genesenen schwerwiegende *Hirnschäden* verursachen kann

‘British neurologists have now published shocking details in the journal "Brain," according to which SARS-CoV-2 can cause severe brain damage even in patients with slight symptoms or already recovered people.’

This does not explain *Hirnschaden*, but it establishes a shocking relation with SARS-CoV-2.

It is surprising and seems particularly worth stressing that coronavirus-related examples in the German Wiktionary are even given to illustrate everyday lemmas, such as *sondern* ‘but’:

(10) Mittlerweile gibt es zahlreiche Hinweise, dass das neuartige Coronavirus SARS-CoV-2 nicht nur die Lungen und Atemwege, *sondern* auch andere Organe massiv angreift.

‘Meanwhile there are numerous hints that the novel coronavirus SARS-CoV-2 massively attacks not only lung and respiratory system, but also other organs.’

Both quotes are formally correct. Whether the assumption of the study is correct, is something else. As a matter of fact, the article on the same study in the aforementioned medical journal *Deutsches Ärzteblatt* (URL28) notes that in none of the patients could the virus be found in the brain and that it is unclear how frequent such damages are. This is not discussed and cannot be discussed in a lexicographic entry, but lexicographic entries should refrain from using such quotes. It seems as if User-de2’s aim is to paint a threatening picture of the coronavirus.

User-de2 also does this with other everyday words that are well entrenched in existing reference books. For de. *grob* ‘gross, large,’ the user adds a quote on a coronavirus study that has become quite famous in the German media. Due to being clipped from the rest of the article, from the original study, and isolated from other comments on the same study, the quote makes the reader think that the study is worthless:

(11) Die Heinsberg-Studie zur Dunkelziffer der Corona-Infektionen basiert auf einer *grob* falschen Berechnung.

‘The Heinsberg study on the estimated number of unreported cases of corona infection is based on a grossly false calculation.’

Actually, this is the subheading in the original source, whose message is corrected in the source itself (URL29). In the entry *de. Suboptimal*, another adjective that has no specific medical sense, User-de2 added the following example on 12 July 2020 and supplied *Coronaerkrankungen* ‘corona diseases’ in square brackets, because the quote does not even include the word *corona* – an addition unnecessary for the illustration of the lemma’s meaning, though:

(12) [Coronaerkrankungen:] "[...] Alles, was der Lunge schadet, ist eher *suboptimal*."

‘[Corona affections:] "[...] Everything that harms the lungs is rather suboptimal.’

User-de2 shows similar behavior with other everyday words, as well; and even if administrators also view and contribute to these articles, they do not criticize this. One could suspect conscious propaganda or fear-mongering. An exception is User-de2’s insertion of the following quote under the lemma *viel* ‘much’ (employing the superlative form):

(13) Die *meisten* Menschen, die sich mit dem neuen Coronavirus SARS-CoV-2 infizieren, erkranken nicht schwer.

‘Most people getting infected with the new coronavirus SARS-CoV-2 don’t fall seriously ill.’

This one instance, however, could also serve as a neutrality fig leaf veiling his general editing behavior. At any rate, this shows a general problem, when a very limited number of maybe well-meaning people, in the sense of simply aiming to enlarge Wiktionary, contribute in a certain area, but are not reminded by others of the Wiktionary rules: Wiktionary may end up with political, potentially propagandistic, but certainly non-lexicographic elements – even if these are mainstream views or mainstream ideologies.

In French entries, too, coronavirus-related quotations are not infrequently found with lemmas consisting of metaphorical idiomatic expressions and even proverbs, without an apparent connection to the coronavirus theme. Are these attempts then to promote certain ideologies?

(14) [s.v. *gros bras* ‘robust person’:] La crise du coronavirus montre que le régime économique et social actuel est mortifère

‘The coronavirus crisis shows that the current economic and social regime brings death’

(15) [s.v. *au grand jour* ‘publicly’:] Face au coronavirus, une Amérique aux pieds d’argile

‘Facing coronavirus, an America built on sand’

In these and three other entries, the attached source title could be interpreted as unnecessarily promoting a certain ideology. These cases may seem negligible, were it not that all five were amended by the same user (we will call him User-fr1). Is someone trying to influence Wiktionary users’ views in a certain direction? Should the user not get comments that several

source titles are highly emotional and could thus negatively affect certain users? This is perhaps not done because User-fr1 has been bestowed with the title "patrouilleur" (literally "patrolman"), an indication that s/he is trusted by the community. Moreover, User-fr1 is also behind the additions of neutral quotes and source titles in these same five lemmas. Again, it seems that an eager contributor is driven to enlarge Wiktionary and thus lacks a certain sense of problematic entry aspects.

Highly connotative coronavirus-related quotes under non-coronavirus-related lemmas are rare in English (e.g., *go by, last straw*). In Dutch, they do occur occasionally in a specific way, which will be treated below. Let us first analyze the use of discriminating certain countries or regions through the selection of examples. Some geographical names are related to the coronavirus topic. In de. *Wuhan*:

(16) Es ist Zeit, dass die Welt die Wahrheit über die Entstehung von Sars-CoV-2 in *Wuhan* erfährt.

‘It is time that the world is told the truth about the evolution of SARS-CoV-2 in Wuhan.’

It is certainly true that the toponym, or place name, *Wuhan* would arouse the connotation ‘SARS-CoV-2’ in Germans’ minds when this was written; however, this specific use also activates the connotation ‘lie, untruth.’ Whether this is a prominent connotation of *Wuhan* should first be investigated in an appropriate scientific analysis. Otherwise, it sounds like anti-Wuhan or anti-Chinese propaganda. The lemma en. *wake-up call* contains the following discriminating quotation (without explicitly mentioning China):

(17) The coronavirus could be a *wake-up call* for people to fully understand the immediate danger of communism.

An additional relevant geographical entry is en. *East Turkistan*:

(18) The Uyghur Human Rights Project is deeply concerned by alarming new evidence that local residents are starving across the Uyghur homeland. Since mid-February, Uyghur-language social media has lit up with disturbing videos, photos, and other information providing evidence that a month-long coronavirus (COVID-19) lockdown is leaving many East Turkistan residents hungry.

This may be an indirect criticism against lockdowns. In the French version, there are many toponyms connected with the coronavirus topic: *Hubei, Anhui, Nigéria, Brazzaville, Deir al-Asad*. In the Polish version, five lemmas of toponymic adjectives include the same quote which mentions the origins of 18 coronavirus-infected people: *jarociński, kępiński, ostrzeszowski, pilski, śremski*.

Sometimes the geographical discrimination effect is part of a quote in a non-toponymic lemma. "China" appears, for instance, in nl. *infectiegeval* ‘infection case,’ nl. *verkenningstocht* ‘expedition,’ fr. *virologiquement* ‘virologically,’ and fr. *fruit sec* ‘dried fruit.’ "South Korea" is part of the example sentences in nl. *superverspreider* ‘superspreader’ and the synonymous *superbesmetter* – these are not citations, just a sentence made up by a user (we will employ the pseudonym User-nl1), actually the same sentence for both words:

(19) De {superverspreider/superbesmetter} in Zuid-Korea zorgde ervoor dat 60%-80% van de ongeveer 8000 besmette mensen met het coronavirus in het land tot één bron te herleiden zijn.

‘The superspreader in South Korea therefore caused that 60-80% of the approximately 8,000 people infected with coronavirus in the country can be traced back to one single source.’

Nord-Brabant is depicted as *infectiehaard* ‘focus of infection,’ with many a *besmettingsgeval* ‘infection case’ in sentences again constructed by User-nl1. Constructing example sentences instead of using already documented sentences is in line with the Dutch Wiktionary regulations only for sentences of up to five words, but it may be asked whether it is a wise decision to construct sentences with assumptions that include unascertained figures. Such sentences will be prone to harm a certain group of people.

Occasionally, personal names are strangely connected to the coronavirus. In the Russian Wiktionary, the entries for two family names (*Гайдук Gayduk*, *Васильев Vasil'ev*) include quotations in which non-VIPs with COVID-19 symptoms who accidentally bear the family name are mentioned.

For the names of concrete persons, let us consider the French version, whose rules include to avoid all quotes transmitting ideologies. While prominent critics of mainstream views around the coronavirus topic do not appear in German quotes, this is different in the French Wiktionary. The most prominent critic of French government measures was already mentioned above, Didier Raoult, whose name led to the coinage of the words fr. *raoulien* and *raoultiste*. In the lemma fr. *antidote*, his name is embedded in a critical, but not too critical context. Raoult is also mentioned in a quote added by a user we will call User-fr2 on 16 May 2020 in the lemma fr. *grippette* ‘little flu’:

(20) Depuis le mois de décembre, même si les gouvernements voulaient se voiler la face et disaient qu'on avait affaire à une petite *grippette* qui n'aurait pas beaucoup de conséquences, je savais par Didier Raoult que, malheureusement, cette vague submergerait la planète tout entière.

‘Since the month of December, even if the governments wanted to turn their heads away and were saying that this was a little flu that didn't have many consequences, I knew from Didier Raoult that unfortunately that wave would flood the whole planet.’

The quote definitely shows the meaning well, due to its contrast of *grippette* and *vague* ‘wave.’ But even though two views are presented, it highlights the latter view. Although no coronavirus-related word is used in this quote and although it occurs under the sub-sense ‘barely virulent flu virus,’ the text undoubtedly refers to SARS-CoV-2 (URL30). It is interesting to note that the next sub-sense of *grippette* is ‘COVID-19’ with the stylistic label “ironic.” Here we have this quote added on 8 July 2020 by User-fr2 again:

(21) Coronavirus : au Brésil, le président, Jair Bolsonaro, testé positif à « la *grippette* »

‘Coronavirus: in Brazil, the president, Jair Bolsonaro, tested positive for "the little flu"’.

Quotes 20 and 21 are on the same topic – with an interesting difference: In quote 20, the fact that political leaders, including the French president, termed coronavirus a small flu is not made explicit; in quote 21, it is not the French but the Brazilian president who used the term,

and there seems to be no problem in revealing that it is about the coronavirus, as if the COVID-19 classification as ‘little flu’ can only be ridiculed if made by the Brazilian president. In quote 20, the Wiktionary user could have added another sentence from the original source which would have included *Covid-19*.

In nl. *epidemioloog* ‘epidemiologist,’ *architect* ‘architect,’ *omstreden* ‘controversial,’ and *beteugelen* ‘to rein in’ the same quote criticizes Sweden’s strategy, which was suggested by their chief epidemiologist (his name is not explicitly mentioned):

(22) De epidemioloog en architect achter de omstreden Zweedse corona-aanpak zegt dat er in zijn land meer gedaan had moeten worden om het virus aan het begin van de uitbraak te beteugelen.

‘The epidemiologist and architect behind the controversial Swedish corona approach says that more should have been done in his country to rein in the virus at the beginning of the outbreak.’

The metaphorical meaning of *architect* is definitely not new, but it was only added to the Dutch Wiktionary entry in 2020, with this one quote above and the incorrect labeling ‘profession.’ Another anti-Swedish quote was appended to nl. *verzorgingshuis* ‘nursing home,’ *naburig* ‘neighboring,’ and *persbureau* ‘press agency.’ All these Sweden-related citations were added by the same user (we will call him User-nl2). Again, it gives the impression that this user wants to promote a view against the Swedish strategy. However, one more citation selected by him occurs in nl. *horecazaak* ‘HORECA business [hotel, restaurant, café],’ but it is neutral. Moreover, User-nl2 even uses one and the same quote that casts Sweden’s chief epidemiologist, Anders Tegnell, in a rather positive light in the lemmas *coronaregime* ‘catalog of corona rules,’ *aanpak* ‘approach,’ *advies* ‘advice,’ *controversieel* ‘controversial,’ and *lockdown*. Although the guidelines say that maximally two quotes should be used, the lemma nl. *lockdown* includes three plus two quotes in two sub-entries that show at least an awareness that official usage and everyday usage differ:

"(regering) officiële instructie waardoor mensen of informatie een plaats of gebied niet mogen verlaten of binnenkomen"

‘(government) official instruction by which people or information must not leave or enter a place or area’

and

"periode waarin mensen thuis moeten blijven om verspreiding van een besmettelijke ziekte tegen te gaan"

‘period in which people must stay at home to confront the spreading of an infectious disease.’

A considerable number of quotes cast US president Trump in a negative light, e.g.:

(23) [s.v. *compound* (verb):] This latest example of nationalistic self-interest **compounded** anger across the EU over Trump’s travel ban, imposed last month without consultation or scientific justification.

Other English anti-Trump quotes occur in *badge of honor*, *budget hawk*, *tick tock*, *Sharpiegate*, and *ruination*, all drawn from the same article entitled "US's global reputation

hits rock-bottom over Trump's coronavirus response.” In the German entry *prüfen* ‘examine,’ User-de2 quotes:

(24) US-Präsident Donald Trump will Forscher prüfen lassen, ob man Corona-Patienten Desinfektionsmittel spritzen kann. [...] Seine Beraterin saß daneben und wirkte unglücklich.

‘US president Donald Trump wants to have researchers examine whether you can inject disinfectants into Corona patients. [...] His advisor was sitting next to him and seemed unhappy.’

Other German entries with Trump-critical example sentences are *allen Ernstes* ‘in all seriousness’ and *im Zusammenhang mit* ‘in connection with.’ French lemmas with Trump-critical quotes are *gourmander* ‘to chide,’ *panne* ‘breakdown,’ and *tailler* ‘to cut.’ Finally, Trump is made fun of in an alternative way under fr. *poisson d’avril*:

(25) Ce matin, j’ai eu l’idée de cueillir ma petite sœur au réveil par un "Tu sais pas quoi? Trump est mort du corona!" [...], à quoi j’aurais ajouté un tonitruant: "Poisson d’avriiiiil!"

‘This morning, I had the idea of waking my little sister up by saying "You know what? Trump died of corona!" [...] after which I would have added in a thundering voice "April Fooools!"’

Cultural Characteristics

Beyond the formal conventions and preferences delineated in the section “Promoting Propaganda,” a few more patterns can be observed. This is done with all due care, since some individual users are especially prolific and their choices cannot automatically be generalized to the entire Wiktionary community.

In the English Wiktionary, there seems to be – apart from anti-Trump quotations – a general predilection to choose quotes with personal names, irrespective of their name recognition, from UK Prime Minister Johnson under *televised* to "Dan Levin, who runs a small company outside Chicago" under *playbook*. Other lemmas with such quotes are *stone-faced*, *glue*, and *nurse*.

In German, among those contributing to coronavirus-related items, there seems to be a preference for direct or indirect quotations within quotations – either by persons or by organizations. Examples 4 and 5 already illustrated this. It is also in the German Wiktionary that we find the most explanatory notes in square brackets. These notes, as already shown, are predominantly used to indicate that the quote is in relation to the coronavirus topic.

In French, time and again, people in political power frame crises rhetorically as war. The conceptual metaphor CRISIS IS WAR has also shaped the discourse on the current coronavirus topic, for instance by France’s president Macron in his speech to the nation on 16 March 2020 (URL31). The war metaphor is also found in English and Spanish, as Silvera-Roig has illustrated in a paper available online, with some examples prominently criticized in the press (URL32). In France, in contrast, this metaphor seems more accepted. Similarly, the French Wiktionary includes "war" citations in coronavirus-related lemmas, e.g., under *Lopburi*:

(26) Les macaques de la ville de **Lopburi**, dans le centre de la Thaïlande, sont sur le sentier de la guerre, et c'est à cause du coronavirus

'The macaques in the town of Lopburi, in the center Thailand, are on the path of war, and it's because of the coronavirus.'

In addition, we find coronavirus quotes in "war" lemmas such as *guerre sanitaire* 'health war' (and also in the semantic equivalent nl. *gezondheitsoorlog*, with reference to the French president).

While a certain humorous creativity is reflected in some of the French neologisms listed in the sections on innovations with and without coronavirus-related morphemes, examples sentences expressing a sense of humor are rare. Here is one from the lemma *mandarin* 'Mandarin, here: literate, educated' – a word originally denoting a high-ranking official in the former Chinese imperial civil service, but connoting in this example a virus seen as coming from China:

(27) le système marque aussi trop de révérence envers les *mandarins* de la science.
« Nous croyons trop, nous ne vérifions pas assez, c'est ce que cet épisode doit nous enseigner. »

'the system also shows too much awe for the Mandarins of science. "We believe too much, we do not check enough, that's what this episode teaches us."'

Synopsis

In a nutshell, coronavirus-related quotations occur even with lexemes of general language that are not specifically semantically related to the coronavirus. Furthermore, there are violations of guidelines, specifically concerning neutrality. This happens across languages, predominantly in German, French, Dutch, and English. Moreover, there are language-specific predilections in the selection of examples.

Discussion, Interpretation, and Conclusion

The coronavirus-related lemmas in the Wiktionary versions in Danish, Dutch, English, French, German, Hungarian, Italian, Norwegian, Polish, Russian, Spanish, and Swedish are presumably provided mostly by untrained lexicographic contributors. Although the analysis was not strictly quantitative, it nevertheless enables a few contrastive characterizations:

1. Concerning RQ1, neologisms have been lemmatized predominantly in the French, English, and Dutch versions, and to a smaller degree in German and Russian. In French and Dutch, this even includes nonce words, whereas a traditional dictionary would exclude these.
2. As to RQ2, it can be said that definitions include coronavirus-related words only if they are truly necessary; they do not occur in superordinate terms such as *pandemic*.
3. Regarding RQ3, it can be stated that guidelines for definitions are not always met. There are inadequate or incomplete definitions, and there is inadequate labeling of words. Definitions sometimes miss that a certain lexeme is used for a health-related concept but is not part of the medical jargon (the label "medical" or "medicine" is inappropriate here) or that the lexeme or one of its variants has a different definition in official medical jargon. Particularly in French, definitions sometimes run the risk of discriminating (especially if

co-hyponyms are treated differently). Discrimination is even committed by high-status contributors.

4. RQ4 must also be answered in the affirmative: Coronavirus-related words do noticeably occur in quotes even under lemmas not related to coronavirus.
5. With respect to RQ5, it can be said that counter to Wiktionary guidelines, example sentences are not always good illustrations of a lemma's meaning, even when added by high-status contributors. Moreover, unexpectedly, coronavirus-related quotations are even added under headwords that are not specifically semantically related to the COVID-19 issue. Further, counter to the guidelines, particularly in German, French, Dutch, and English, example sentences partially purvey propagandistic views, in the sense that they reflect certain stereotypes, ideologies, or discriminations – even those provided by high-status contributors. This also occurs with words and names unrelated to the coronavirus topic – particularly in German and French, and occasionally in English, Polish, and Russian.
6. An additional observation is that in the Dutch version – contrary to the guidelines – example sentences are rarely connected to a source.
7. Another additional observation is that the inclusion of proper names in example sentences appears more typical in the English version than in the other language versions. In German, example sentences include more (direct or indirect) quotations from people or organizations and include explanatory notes in square brackets more often than the other versions. In the French version, they display conspicuously more war metaphors than in the other versions.

Of course, there are limits to generalizability, since only one societally "hot" topic (six months after its emergence) was investigated, and other topics – including "hot" topics – may attract different contributors. Nonetheless, points 3 to 6 show that Wiktionary still needs improvement in quality management, not only with definitions, which Fierres Olivera (2009) previously highlighted, but also with examples. This relates especially to the French, Dutch, German, and English communities – in other words, the western languages, including German, whose Wikipedia version was already presented in the literature review section as lacking neutrality. This also supports our theory that in Wiktionary it can take time to benefit from collective intelligence. The reason may be because Wiktionary users do not typically refer to Wiktionary to check examples and thus fewer users than in Wikipedia note the problems. However, definitions also turned out to be problematic. Whether the violations are conscious view-shaping efforts or whether the contributors were driven by other motives (e.g., overcoming their own COVID-related stress) cannot be determined from this analysis.

From a lexicographic perspective, it may be justifiable to use example sentences not only to illustrate the meaning of a word but to reflect stereotypical contexts of use. However, this should be done with care when it concerns words from ongoing, highly emotional topics where majority attitudes can shift. Moreover, it should be made clear when semantic descriptions are of connotation rather than denotation. It cannot be justified from a lexicographic point of view, though, to invoke stereotyping and discriminating contexts with keywords from outside the specific emotional topic. Wiktionary users, too, should be made more aware of this issue.

Notwithstanding this criticism, Wiktionary remains a valuable first source for cross-linguistic synchronic and diachronic information and for corresponding links to further reading. For this, all contributors should be thanked.

References

- Bronner, G. (2013). *La démocratie des crédules*. Paris: PUF.
- Callahan, E. S., & Herring, S. C. (2011). Cultural bias in Wikipedia content on famous persons. *Journal of the American Society for Information Science and Technology*, 62(10), 1899-1915.
- Doshi, P. (2011). The elusive definition of pandemic influenza. *Bulletin of the World Health Organization*, 89, 532-538. <https://www.who.int/bulletin/volumes/89/7/11-086173/en/>
- Duhamel, O., & Bigorgne, L. (2020). *Les mots du coronavirus*. Paris: Dalloz.
- Fuertes-Olivera, P. A. (2009). The function theory of lexicography and electronic dictionaries. In H. Bergenholtz, S. Nielsen, & S. Tarp (Eds.), *Lexicography at a crossroads* (pp. 99-134). Bern: Lang.
- Gouws, R. H., Heid, U., Schweickard, W., & Wiegand, H. E., Eds. (2013). *Dictionaries – Supplementary volume*. Berlin: Walter de Gruyter.
- Gredel, E. (2018). Wikipedia als Reflexionsgegenstand in sprach- und mediendidaktischen Kontexten. In: M. Beißwenger & M. Knopp (Eds.), *Soziale Medien in Schule und Hochschule* (pp. 165-189). Frankfurt: Lang.
- Greenstein, S., & Zhu, F. (2018). Do experts or collective intelligence write with more bias? *MIS Quarterly*, 42 (3), 945–959.
- Grzega, J. (2019a). *Wort-Waffen abschaffen*. Berlin: epubli.
- Grzega, J. (2019b). Names, nations and nature from a Eurolinguistic view. *Journal for Eurolinguistics*, 16, 23-36. <http://www.eurolinguistix.com>
- Harras, G. (1989). Zu einer Theorie des lexikographischen Beispiels. In: F. J. Hausmann et al. (Eds.), *Wörterbücher/Dictionaries/Dictionnaires*, vol. 1 (pp. 607-614). Berlin/New York: de Gruyter.
- Hausmann, F. J. (1989). Die Markierung im allgemeinen einsprachigen Wörterbuch.: In: F. J. Hausmann et al. (Eds.), *Wörterbücher/Dictionaries/Dictionnaires*, vol. 1 (pp. 649-657). Berlin/New York: de Gruyter.
- Hausmann, F. J., Reichmann, O., Wiegand, H. E., & Zgusta, L., Eds. (1989). *Wörterbücher/Dictionaries/Dictionnaires*. Vol. 1. Berlin/New York: de Gruyter.
- Hoeres, P. (2015). Hierarchie in der Schwarmintelligenz. In: T. Wozniak, J. Nemitz, & U. Rohwedder (Eds.), *Wikipedia und Geschichtswissenschaft* (pp. 15-32). Berlin: de Gruyter Oldenbourg.
- Hwang, T., Bourgeois, F. T., & Seeger, J. D. (2014). Drug safety in the digital age. *New England Journal of Medicine*, 370(26), 2460–2462.
- Martin, R. (1989). L'exemple lexicographique dans le dictionnaire monolingue. In: F. J. Hausmann et al. (Eds.), *Wörterbücher/Dictionaries/Dictionnaires*, vol. 1 (pp. 599-607). Berlin/New York: de Gruyter.
- Mehler, A., Gleim, R., Hemati, W., & Uslu, T. (2018). Skalenfreie online-soziale Lexika am Beispiel von Wiktionary. In: S. Engelberg, H. Lobin, K. Steyer, & S. Wolfer (Eds.), *Wortschätze* (pp. 269-292). Berlin: Walter de Gruyter.

- Pfeil, U., Zaphiris, P., & Ang, C. S. (2006). Cultural differences in collaborative authoring of Wikipedia. *Journal of Computer-Mediated Communication*, 12(1), 88-113. <https://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2006.00316.x>
- Reiss, K., & Bhakdi, S. (2020). *Corona: False alarm?* Chelsea: Chelsea Green.
- Wagner, C., Garcia, D., Jadidi, M., & Strohmaier M. (2015). It's a man's Wikipedia? Assessing gender inequality in an online encyclopedia. *Proceedings of the International AAAI Conference on Web and Social Media (CWSM)* 9(1), 454-463. <https://ojs.aaai.org/index.php/ICWSM/article/view/14628>

URLs (all retrieved 19 August 2020)

- URL01. https://en.wikipedia.org/wiki/Wikipedia:Neutral_point_of_view
- URL02. <https://de.wiktionary.org/wiki/Wiktionary:Standardreferenz>
- URL03. <https://de.wiktionary.org/wiki/Wiktionary:F%C3%BCnf-Zitate-Regel>
- URL04. <https://de.wiktionary.org/wiki/Hilfe:Zitate>
- URL05. <https://nl.wiktionary.org/wiki/WikiWoordenboek:Definities>
- URL06. <https://en.wiktionary.org/wiki/Wiktionary:Quotations>
- URL07. https://en.wiktionary.org/wiki/Wiktionary:Example_sentences
- URL08. <https://fr.wiktionary.org/wiki/Wiktionnaire:Exemples>
- URL09. <https://www.who.int/mediacentre/news/notes/2015/naming-new-diseases/en/>
- URL10. <https://www.termcat.cat/ca/diccionaris-en-linia/286/presentacio>
- URL11. <http://anw.inl.nl/search>
- URL12. <https://www.taalkbank.nl/2020/03/14/coronawoordenboek/>
- URL13. <https://public.oed.com/blog/>
- URL14. https://www.oxfordlearnersdictionaries.com/wordlist/new_words
- URL15. <https://www.germanistik.uni-wuerzburg.de/sprawi/forschung/seuche-und-sprache/forschung-aktuell/>
- URL16. <https://www.dwds.de/themenglossar/Corona>
- URL17. http://www.treccani.it/magazine/lingua_italiana/autori/Pietrini_Daniela.html
- URL18. https://verne.elpais.com/verne/2020/04/07/articulo/1586246728_179666.html
- URL19. <https://www.rae.es/noticias/crisis-del-covid-19-sobre-la-escritura-de-coronavirus>
- URL20. <https://de.wiktionary.org/wiki/Verzeichnis:Deutsch/Corona>
- URL21. <https://www.rivm.nl/coronavirus-covid-19/grafieken>
- URL22. https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Fallzahlen.html
- URL23. https://info.gesundheitsministerium.at/dashboard_Epidem.html?l=de
- URL24. <https://www.bag.admin.ch/bag/de/home/krankheiten/ausbrueche-epidemien-pandemien/aktuelle-ausbrueche-epidemien/novel-cov/situation-schweiz-und-international.html>
- URL25. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/>
- URL26. <https://de.wiktionary.org/wiki/Wiktionary:Literaturliste/Fachliteratur>
- URL27. <https://www.aerzteblatt.de/archiv/214370/PCR-Tests-auf-SARS-CoV-2-Ergebnisse-richtig-interpretieren>
- URL28. <https://www.aerzteblatt.de/nachrichten/114536/SARS-CoV-2-Studie-beschreibt-breites-Spektrum-neurologischer-Komplikationen>
- URL29. https://www.tagesschau.de/investigativ/swr/heinsberg-studie-103.html?utm_source=pocket-newta
- URL30. https://www.liberation.fr/politiques/2020/04/14/les-intox-d-estrosi-pour-defendre-didier-raoult_1785142

URL31. <https://www.elysee.fr/emmanuel-macron/2020/03/16/adresse-aux-francais-covid19>

URL32. https://www.researchgate.net/publication/341881753_Conceptualizing_Covid-19_global_war_against_an_invisible_enemy

Biographical Note

Joachim Grzega [joachim.grzega@ku.de] is Associate Professor of Linguistics at the University of Eichstätt-Ingolstadt and Head of the Unit “Innovative Europäische Sprachlehre (InES)” (Innovative European Language Teaching) at VHS Donauwörth. His current research interests are contrastive linguistics and Eurolinguistics (historically and synchronically, with respect to lexicology and pragmatics/discourse); the connection between language, thought, and culture (including political topics); efficient language teaching (especially at the basic level); intercultural communication (especially English as a lingua franca); and expert-layperson communication (especially in edutainment).