Francesco Meledandri

The Impact of Polarised Social Media Networking Communications in the #Longcovid Debate between Ideologies and Scientific Facts

A Language Analysis from a Corpus of Posts on Twitter/X

Abstract

The Covid-19 pandemic has shaped many conventions in social, cultural and linguistic terms. Its outbreak in early 2020 sparked an unprecedented R&D effort (Pedrini 2021; Burgess et al. 2021) by the international scientific community in developing effective vaccines (Knoll and Wonodi 2020). Indeed, these vaccines helped the world to return to normal life, though the vaccines' effectiveness was questioned (Andrews et al. 2022).

The global debate concerning Covid-19 issues has not come to an end, especially in online and digital environments. Media play roles in creating and shaping representation(s) and truth(s) (Garfin et al. 2020; Chaiuk and Dunaievska 2020), but it is on social media networking platforms that discussions have been leading to patterns of polarisation. A case in point is Long Covid, or post-Covid-19 condition, defined "as a variety of mid- and long-term effects after [people] recover from their initial illness" (WHO 2021). Online discussion about this topic has increased over time.

The aim of this study is to present the results of an analysis of a corpus of automatically-retrieved (Brooker, Barnett and Cribbin 2016) social media networking content from Twitter (now called "X") in a 13-week timespan (September 2022–November 2022) and made up of almost 600,000 tweets. The 2.5M-token corpus of tweets is investigated using quantitative and qualitative approaches (Stefanowitsch 2020) to retrieve data on users' reactions about Long Covid. Some of these statements reflect conspiracy-based theories involving vaccination, fake news and post-truths, clashing with scientific evidence. Other tweets reflect supportive stances, thus leading to forms of ambient affiliation (Zappavigna 2011).

Keywords: *Covid-19, discourse of pandemics, Long Covid, social media networking interactions, corpus linguistics*

1. Introduction

Covid-19, which first emerged in late 2019 and had evolved into a pandemic by March 2020 (Cucinotta and Vanelli 2020), caused a seismic shift in the way we live. The pandemic's peak was in 2020-2022 in terms of casualties and coping strategies. Even today it poses unprecedented challenges in medical terms and has changed everyday habits, protocols and routines. The pandemic made the entire global community reconsider conventional schemes in social and cultural terms, causing new paradigms to be enacted and scrupulously followed to save lives (UN 2021).

Such paradigms reflected the efforts of the scientific community to deal with this unexpected emergency. From 2020 on, Covid-19 as a global topic has been characterised by several different stages and methods for managing the "invisible enemy" (Shaw 2020, 531). The first stage corresponded to the sudden outbreak of the virus, which was battled by means of forced isolation and related protocols. The second stage saw the invaluable effort of R&D worldwide aimed at the development of effective vaccines against the virus. Within a few months of the beginning of the pandemic, the vaccine *race* (Pedrini 2021; Burgess et al. 2021) started. It led to the massive production of sera yielding encouraging results in the fight against the disease (Knoll and Wonodi 2020; Voysey et al. 2020; Polack et al. 2020). Vaccines have helped people from a direct medical point of view, but they have also helped indirectly by restoring pre-pandemic conventions and habits (Li and Giabbanelli 2021). As more people were vaccinated, large-scale *in vero* results proved to be less effective than expected, diminishing the energy of the initial thrust, even from an emotional point of view (Liu and Chu 2022). At the same time, hesitation about being vaccinated proved to be a decisive factor impeding global immunization (Sussman et al. 2023; Al-Amer et al. 2022). The emergence of virus variants led to successive surges in infections and deaths, and even though vaccines did protect people, their reduced *ex-post* effectiveness (Andrews et al. 2022; Katella 2022) opened a new stage of debate about them. From 2022 on, a complicated coexistence with the virus has become the norm. This phase is still ongoing even though the end of the Covid-19 health emergency was declared on May 5, 2023 (WHO 2023).

Three years after the beginning of the pandemic, Covid-19 is still part of our lives even though it is much more easily managed. Even so, its aftermath is still to be assessed in tangible terms due to its several repercussions in economic terms, but more importantly in social, cultural and psychological terms (Saladino, Algeri and Auriemma 2020; Nelson 2020). Long Covid is a case in point. Long Covid is a post-Covid-19 condition that has been officially recognised "regardless

of age or severity of original symptoms" (WHO 2022) and which has affected 10-20% of people infected by the virus. As the expression suggests, Long Covid is:

> The continuation or development of new symptoms 3 months after the initial SARS-CoV-2 infection, with these symptoms lasting for at least 2 months with no other explanation, […] that can have an impact on everyday functioning. (WHO 2022)

With more than 65 million cases recorded as of this writing, more than 200 symptoms have been attributed to this medical condition (Davis et al. 2023). Clinical matters such as issues with testing procedures have contributed to cast Long Covid in a bad light; at the same time, "inaccurate pandemic narratives and widespread lack of postviral knowledge have caused downstream issues and biases in long Covid research and care" (Davis et al. 2023, 141). This condition deserves in-depth analysis, particularly because for lay people, Long Covid represents both a medical condition and a battlefield of language and communication.

This paper aims to provide a language analysis of a collection of texts representing a sample of actual communications dealing with Long Covid. The collection of texts is taken from a popular and free social media platform (Twitter, now rebranded X) used in digital communication to create conversation chains and contact networks, thus sustaining a proper digital space where affiliation is created with discourse-oriented aims (Zappavigna 2012). User-generated content is written/posted for a specific purpose and aimed at creating communities of users gathered around language-oriented but also ideological concerns (Hinton and Hjort 2013, 60). Thus, the purpose of this study is to observe the kinds of discourse arising from "free" spaces where ordinary users voice their opinions with different levels of expertise in a particular field. Notwithstanding gaps in terms of knowledge, some messages involving a specialised field (Long Covid, in this case) may achieve considerable resonance when conveyed via digital networking platforms. A corpus-based analysis of the language sample collected for this purpose will provide an interpretation of the communication acts concerning users' attitudes when facing Long Covid. This condition is voiced from both an external perspective (as a condition to talk about) and a personal stance due to its widespread diffusion. This is so even though its aetiology is not objectively clear, thus causing a mixed representation of facts within the same networking platform. Therefore, this study observes how the language found in social media networking platforms covers specialised topics, causing the alteration of factuality. The hypothesis is that when found on social networking platforms, Long Covid-related communication is discussed with polarised (mostly negative) views, but instances of neutral—say, scientific—

communication acts may be found as well, since some people and institutions might prefer sharing facts rather than personal opinions.

2. Discourse on social media platforms: a catalyst for (post)truths and (anti)facts

The advent of information technology led to innovative communication in many forms. Categories such as time and space have been revolutionised thanks to contemporary digital and/or online interactions, "loosening the bond between activities and fixed location and times" (Ben-Elia and Zhen 2018, 268). Real-time activities with no physical hindrance have become common, thus broadening the reach of such interactions in terms of intended audience and persistence over time. These innovations led to the development of communities of people—users, aggregating according to the digital channels they use. People tend to gather around common interests, geographical origin or language, enacting distinctive communication and discourse practices. These online communities have been around for almost thirty years, "transforming the socio-economic structure of work and offering new environments for daily life" (Bainbridge 2020, 1). The perfect sublimation for such practices is represented by Social Networking Sites (SNSs), that is free platforms "with a tenacious social structure" (Kapoor et al. 2018, 531) through which users can communicate and interact with (almost) no constraints. The field of Social Media Studies developed quickly (Vittadini 2018), with a related interest in forms of discourse that can be found in any online community, even outside SNS-based digital practices (KhosraviNik 2022). However, social media's impact in creating discourse(s) is indisputable due to the implication they carry in terms of meaning-making processes and the diverse consequences of social and cultural constructs (Tamássy and Géring 2021), but also with repercussions in real life (Damota 2019) in both positive and negative terms.

In the realm of SNSs, forms of communication with no direct connection (i.e., they are not part of the same conversation) could be gathered and retrieved for discourse-related purposes. This is why a language analysis of topics such as Covid-19 or Long Covid can reveal ideological vs. fact-based intentions. SNSs have become digital repositories of personal opinions, and the widespread misuse of such platforms has led to the proliferation of two kinds of language events that are diametrically opposed. On the one hand, there is space for objective, reliable, fact-based evidence advocated by verified stakeholders and experts. On the other hand, the propagation of pseudo-information that circulates freely without filtering leads such views to become as popular as actual facts. In fact, the latter invaded public digital spaces to the point of becoming "comfortable information" for users, allowing the phenomenon of "post-truth" to present an

embedded phenomenon of contemporary digital forms of communication (Cosentino 2020; McIntyre 2018; D'Ancona 2017). It is no surprise that the pandemic has come to represent "the most blatant expression of dangers of the post-truth age, characterised by […] a blurring of the line between opinion and fact" (Schulman 2020).

3. Case study: analysing a corpus of tweets involving the Long Covid debate

In light of the many possibilities offered by social media interactions, it is possible to gather a set of data from SNSs for the purpose of analysing users' behaviour in discourses related to Long Covid, which has become a key interest during the second phase of the pandemic (from 2022 on) and envisaged as an illness since the early stage of the pandemic (Callard and Perego 2021). Due to its features as a SNS, Twitter/X can be exploited for different purposes and by diversified groups of users who voice their thoughts about Long Covid either actively or passively (i.e., having suffered from Long Covid or simply by retrieving information or sharing their views). This practice in turn generates a large dataset of texts that can be used to understand how Long Covid is perceived and felt, in order to provide possible medical counteractions (Dolatabadi et al. 2022). Starting from this perspective, data may be collected to create language corpora to be analysed in terms of recurring words and patterns associated with the topic.

3.1 Method: choice of SNS

The first methodological concern to take into consideration is which SNS was used to gather data. In this case, only one platform has been chosen: Twitter/X. Several reasons motivated this choice. First, Twitter/X favours discourse patterns thanks to the use of device-specific tools, such as mentions and hashtags, which can allow for similar contents to be retrieved even though they belong to stand-alone communication acts (Zappavigna 2015; 2012). Thus, data from Twitter/X "lend itself to the exploration of 'topics'" (Brooker, Barnett and Cribbin 2016, 5). Another reason for the selection of Twitter/X lies in the fact that all accounts are public by default, therefore almost all tweets matching a given query can be retrieved. Unlike other SNSs that spur users to use their identity (such as Facebook, where users are more likely to create networks comprised of friends and relatives from *real* life [Georgalou 2017, 256]), many Twitter/X users use pseudonyms and non-identity-based references as their account names since those networks are more likely to be built on content-based or discourse-based interests.

3.2 Retrieval tool: Chorus TweetCatcher Desktop

To gather data from Twitter/X, a specific software was used. Chorus TweetCatcher Desktop (CTD) is a free Social Media analytics tool (Brooker, Barnett and Cribbin 2016). This tool is used to retrieve Twitter/X data according to simple or advanced queries. The software searches for tweets containing hashtags and/or semantically dense keywords related to any subject, in this case Long Covid. In this way, an almost real-time retrieval of content from the Twittersphere is made possible, thus allowing a pragmatic analysis of the data collected (Hoffmann and Bublitz 2017). The retrieval process leads to the creation of social media data that can be analysed as a corpus.

3.3 Timespan and corpus collection

Twitter/X allows for limited automatic retrieval in terms of time (i.e., the last 7 days). To bypass that limitation and collect a corpus of tweets with time-based consistency, a weekly search via TweetCatcher was performed. The period of retrieval extended from September 1, 2022, to November 29, 2022 (13 weeks). Each retrieval session was carried out using the same settings (tweets in English, complete set of metadata). The query used to collect a corpus of tweets was the hashtag #LongCovid, which has been assumed to be the most evocative linguistic element to represent this syndrome since it "gained consistency in just a few weeks" (Callard and Perego 2021). Other studies included other queries such as *long Covid syndrome, long Covid, post-Covid syndrome, post-acute sequelae of SARS-CoV-2, long-term Covid, long haulers, and chronic Covid syndrome* (Awoyemi et al. 2022, 1), but the hashtag alone is sufficient to provide a comprehensive retrieval of SNS-specific discourse practices.

Data obtained from Tweetcatcher were filtered by excluding non-essential metadata. For this purpose, only the actual message sent by the user, that is the Tweet, was included in the corpus. Tweetcatcher saves data in a tab-delimited format. This allows for the isolation of single categories (in this case, the category "Tweet") which may be processed with other applications such as Microsoft Excel or advanced text-editors. The isolation of language-based occurrences allows for the creation of a corpus, with data saved as plain text file and imported into corpus analysis tools. A corpus of SNS-specific elements allows for the investigation of

> Whole datasets as an 'information space' in which semantic features (words, hashtags, etc.) intersect in potentially interesting ways, […with] [r]esearchers […carrying out] the exploration of topical structures emerging from the entire body of data. (Brooker, Barnett and Cribbin 2016, 5)

The corpus-based analysis allows for a deeper quantitative and qualitative investigation of the whole language set (Stefanowitsch 2020). Data are evaluated to reveal users' statements and sentiments about Long Covid. In quantitative terms, the number of tweets collected in the 13-week timespan was 591,852. To minimize noisy data and to eliminate duplicate messages that might alter the interpretation of frequencies in a corpus-based approach, Retweets (a one-click feature in Twitter/X that reposts a user's message in their own timeline, thus broadening the reach of the message) were excluded from the corpus using TweetCatcher's command "Remove Retweets." After this filtering process, the number of unique tweets was 78,083.
The bulk of tweets has been analysed using corpus-processing tools, i.e., AntConc v. 3.4.3 and 3.5.8 (Anthony 2019). In total, the corpus has 2,594,365 tokens and 130,092 word types.

4. Results: a quantitative overview

When analysing discourse-oriented terms, frequency of use provides insight about the nature of the data retrieved, since it represents "a good starting point for the analysis of any type of corpus" (Baker 2006, 47). The most frequently occurring element in the wordlist obtained from this corpus of tweets focusing on Long Covid was *longcovid* (71,679), which is not surprising since it matched the retrieval query (#longcovid). Apart from grammatical items such as articles and prepositions, the top-20 list reveals some unusual elements such as the tokens *t, co,* and *https* (rank 3, 5 and 10, respectively). This result is not surprising, since these are linguistic elements typically found in tweets when a link or multimedia content is included in the communication act. One of the most peculiar features of Twitter/X is its limit of 280 characters per tweet for non-confirmed accounts (representing most users).[1] Thus, multimedia content is typically indicated by a short URL with the domain *t.co*, enclosed in a secure protocol (indicated by *https*). Such elements were considered tool-specific, noisy data which can be left out. A relevant though predictable element in the top-20 list is the token *covid,* with almost 30,000 occurrences (29,154). This result confirms the association between the specific Long Covid syndrome and the illness in general; yet, due to the frequency for Long Covid, it also confirms that the main query has fair independence of use since its standalone frequency is almost 2.5 times greater.

[1] There are only 640,000 formal subscribers on Twitter/X (as of September 2023) compared to 611,000,000 active users (April 2024). See https://www.statista.com/statistics/1389933/number-of-twitter-blue-subscribers/#:~:text=X%20Premium%20allows%20users%20to,640%20thousand%20X%20Premium%20subscribers. And https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/. All websites last visited on 17/06/2024.

Reviewing the frequency list, the number of occurrences drops significantly for other lexical items. Though this may result in weaker semantic connections between the query and its co-text(s) of use, many tokens have a semantic relatedness with the main topic of Long Covid. Some of these tokens belong to the general domain of medicine or refer to the realm of Covid-related conditions, including Health (4,635), Symptoms (4,422), Infection (3,685), Brain (2,617), Pandemic (2,521), Care (2,286), Disease (2,050), Suffering (1,895), Sick (1,812), Illness (1,809), and Damage (1,729). Other terms have a closer connection with Long Covid in particular, such as Vaccine (1,292), Viral (1,231), Pain (1,177), Cause (914), MECFS (7,618), CFS (1,710), Chronic (1,663), Mild (1,175), Fog (1,090), Severe (1,057), Serious (694), Dysautonomia (634), Mental (521), Fibromyalgia (504), and Dysfunction (484). This basic distinction could represent a starting point for a discussion of discourse-oriented elements that characterise the way users address Long Covid.

5. Discussion

As stated above, some elements in the corpus of tweets are domain-specific and are used with specific purposes. A case in point is the token MECFS (7,618), which is found in almost one in 10 tweets in the corpus but belongs to the domain of specialised lexis. The term MECFS is an acronym for Myalgic Encephalomyelitis/Chronic Fatigue Syndrome, defined as a "complicated disorder" (Mayo Clinic 2023) and a "disabling and complex illness" (U.S. Centers for Disease Control and Prevention 2023). Generally speaking, MECFS is used in the corpus as a specialised term since only the acronym is provided (this also may be attributable to the character limit for each tweet), even though collocation with more explicatory hashtags such as *#myalgicencephalomyelitis* (297 hits in a L5-R5 span), *#fybromalgia* (242 hits in a L5-R5 span) and *#chronicfatiguesyndrome* (92 hits in a L5-R5 span) could be observed. As a matter of fact, a list of semantically relevant hashtags with no co-textual continuity (i.e., not being part of a sentence) could be included in a tweet to define and index a certain topic and/or to make sure such communication enhances its searchability within a discourse on a topic (Zappavigna 2015). The term MECFS is used both in information-based interactions such as Twitter/X accounts from medicine portals (which post the results of studies associating MECFS and Long Covid, since the former is a proven symptom of the latter [Wong and Weitzer 2021]), but also as personal experiences voiced by ordinary users with Long Covid describing their symptoms. Some clear patterns arise as terms such as *disabling* are associated with MECFS, especially in tweets with divulgation and information-based purposes ("Mecfs is a disabling disease…"; "LongCovid and MECFS are disabling millions of people…"). In particular, the term *millions* is

relevant to users' attitudes about Long Covid. From a quantitative perspective, the token *millions* is a strong collocate of *MECFS* (439 hits, 89% of them in the 5R span). The occurrence of the two words (*MECFS* and *millions*) in Long Covid tweets/posts shows the reach of this disabling condition by expressing as personal distress (e.g., "a terrifying way of life with #longcovid. #MECFS I'm one of millions that are #notrecovered from Sars-Cov-2"; "I lost my career and life to #mecfs There are #millionsmissing & millions more w/ #longcovid"; "because I was an athlete posts are very useful to the millions with #LongCovid and #Mecfs who weren't athletes") and a collective condition ("Millions of us with #LongCovid, #MECFS, #postvax and #pots are #notrecovered"; "How sad that there are millions of ppl #mecfs #longcovid around the world who would offer themselves…"; "How much talent is lost because millions are sick. #LongCovid #MECFS"; "Long-Covid and MECFS are disabling and causing millions to suffer and kill themselves…"; "#Millionsmissing refers to millions of #mecfs patients missing from their lives"; "We need help with #LongCovid and #MECFS, millions of us are suffering"; "Can you get something done about #LongCovid and #MECFS? Millions of us are suffering"). Other co-texts interrelate these tokens with other topics such as politics. In this case, users post their concerns about possible repercussions of Long Covid on 2022 U.S. midterm elections (which happened during the period of retrieval; see section 3.3) with statements such as "We need help with #LongCovid and #MECFS. How can millions of people vote when bedridden?" More generally, though, connections with politics (especially in U.S.) are voiced using direct tool-specific references: mentions can be enacted by putting a @ before an account's name. Many tweets that raise critical views include names of politicians or their accounts ("#MECFS and #LongCovid are disabling millions of Americans"; "Can you get something done about #LongCovid and #MECFS? Millions of us are suffering"; "We need help with #LongCovid and #MECFS, millions of us are suffering"). These include mentions of @POTUS, @JoeBiden, @KamalaHarris, @RepKClark, just to name a few. The motivation for these mentions is clear. Unlike hashtags, which are used as a discursive element within or beyond the co-text to highlight and gather content (Zappavigna 2015), mentions notify a specific user and publicly indicate who users wish to tell—or even blame—for critical views. In this way, a tweet/post reveals not only the user's intent, but also identifies the relevant institutional stakeholder(s) (the so-called *establishment*, in a derogatory sense when views are negative). This "digital pillory" renders tweets/posts ideological forms of communication, as confirmed by other studies focusing on opposing political groups (Badaan et al. 2023).

Similar to MECFS, the occurrences of *fog* tend to describe a common consequence of Long Covid. In particular, the use of the phrase *brain fog* appears frequently in the corpus (1,000

occurrences), mostly in a sequence of symptoms associated with Long Covid. Co-textual information and collocations for *brain fog* reveal the presence of medical conditions such as *anxiety, breathlessness, chronic/extreme fatigue, cognitive fatigue/impairment, dizziness, dysregulation, distress, headaches, lung issues, memory issues/loss, pain, memory impairment,* and *depression.* The term *fog* evokes a clear symptom. It is used to explain a medical condition in non-technical though explicit terms and is thus suitable for a non-expert community that might stumble upon a tweet describing a symptom they recognise as their own when looking for Long Covid-related information. There are few examples of *brain fog* and MECFS in the same communication act; this confirms that the terms are used interchangeably to define a consequence of Long Covid, functioning as synonyms or defining the same disorder (Azcue et al. 2022). The association between the use of the phrase *brain fog* as a non-technical (hence suggestive) expression and the inner dimension of users experiencing this condition is provided by its collocation with the pronoun *I* (119 occurrences). Most of the examples reveal a negative attitude concerning this undesired condition, so these tweets/posts are framed within forms of digital (self)-narration ("I am mentally exhausted from brain fog, stress"; "I am terrified of brain fog"; "I am having a really hard #LongCovid brain fog and I need help"; "When the brain fog gets bad I get so frustrated"; "the most mentally taxing with my #LongCovid brain fog. I am toast after"; "Between #LongCovid and #Menopause brain fog I am in hell"; "When I say something stupid I blame brain fog"; "That brain fog is real. I am having a hard time"; "Covid brain fog is so frustrating I keep forgetting so many little things"; "I have mild brain fog today"; "I navigate the brain fog of #LongCovid"; "To this day I still suffer from brain fog"; "I am not the person I was. I struggle with fatigue, brain fog"; "Today is a rest day. I'm having really bad brain fog"). In other examples, users give advice, face their condition in supportive terms, and try to bring hope to people who share the same symptoms ("I had an almost overnight improvement in brain fog and PEM when I started taking (magnesium) phosphate"; "you get like aching joints, eyes going funny, dizzy, brain fog and the fatigue. Still I'm working and still alive"; "Still have fatigue and brain fog but I feel improved when I take a hot bath"; "After a #longcovid brain fog crash yesterday, I made it back to work today"; "Beetroot Extract (helps nitric oxide) & B vits to help offset. I've found them all beneficial"; "My #LongCovid brain fog is gone"; "Some things I've noticed: Brain fog substantially better").

As content found on SNSs is often uncontrolled and unfiltered regarding scientific factuality, it can give rise to biased, partisan interpretations of reality. These views are typically found on user-generated content platforms, but they also reflect a tendency of traditional media to emphasize and distort reality/ies concerning the narration of health emergencies (Garfin, Cohen

Silver and Holman 2020; Chaiuk and Dunaievska 2020). A case in point is the controversial relationship between vaccination and (Long) Covid due to the active community of anti-vaxxers and conspiracy theorists who tend to be more persuasive on SNSs than they are in real-life situations (Birchall and Knight 2022). The lemmatised token *vaccin** (5,238 tokens), then, has ideological repercussions in this corpus, too, since a direct association between vaccination and Long Covid is often envisaged. The list of collocations (span: 5L-5R) of this lemma reveals a twofold interpretation of this association. The most frequently used terms represent positive, supportive attitudes through awareness-raising hashtags such as *reducetransmission*, *bringbackmasks* and *covidisnotover*. While less prevalent, tokens such as *infections, vaccineinjuries* and *vaccinesideeffects* are nevertheless 'dangerous' because they spread non-factual information in describing Long Covid. Here, hashtags convey an ideological stance to index the content of a tweet/post and to provide it with a positive or negative connotation. Hashtags have a precise role on Twitter/X and, in discourse terms, they "increase the 'loudness' of their discourse by increasing the likelihood that their words will be found […] In other words, it creates the possibility of *ambient affiliation*" (Zappavigna 2011, 800). An illustrative case is the association between *vaccin** and the collocate *unvaccinated* (51 instances). Almost all concordance lines where these elements appear together tend to provide scientifically based comparisons between vaccinated and unvaccinated people in order to observe the possible correlations of Long Covid, with little or no space for personal views ("But we need to know more clearly the prevalence in unvaccinated (i.e. pre 2021) and vaccinated populations"; "what is the risk of #LongCovid in vaccinated vs. unvaccinated populations? In age groups?"; "It's very difficult to compare vaccinated to unvaccinated ones if they know (or chose) whether they are or not"; "@Independent_ie Vaccinated & unvaccinated people can have #LongCovid"; "[…] found NO DIFFERENCE in long Covid symptoms such as brain fog and fatigue in vaccinated and unvaccinated people. This will trigger some of the pro-vaxxers"; "New #studies have shown that there IS protection from #LongCovid for #vaccinated (10% risk compared to 15% for #unvaccinated)"; "the vaccinated group had only a 19% lower risk of developing #LongCovid than the #unvaccinated group"; "[vaccinated people] had shorter Covid-19 symptom duration and lower risk for #longcovid than unvaccinated"). These illustrations show that tweets, while meant as quick forms of information, often are used wisely and scientifically to spread facts rather than unverified contentions.

Another significant token in this corpus of tweets is *#MedTwitter* (689 hits), which refers to a specific domain for a well-defined audience. This hashtag is placed in tweets/posts dealing with medicine in reference to a community of Twitter users seeking medical advice and networking.

Labelled as an "invaluable resource" by experts in the medical field (Khan 2022), *#MedTwitter* is included in a text to develop a proper medical community and try to provide a sort of certification of the content provided in the tweet. The presence of *#MedTwitter* in this corpus indicates users' desire to include the discourse of Long Covid within this wider stream of medical information, and to make it available for consultation at a later stage. It is no surprise that #MedTwitter appears at the end of a post in which a question or a request to the community has been formulated ("Is there any research out there about #COVID and #LongCOVID causing people to become Protein S deficient and/or causing Protein C decrease or deficiency? #TeamClots #MedTwitter"; "Advice? #MedTwitter #LongCovid #respiratory"; "What is the connection between tinnitus and (small fiber) peripheral neuropathies, if there is any? #MedTwitter #LongCovid"; "Suggestions, tips, advice appreciated! #MedTwitter"; "I felt that weakish sensation that I get some mornings. What could that be? #MedTwitter #LongCovid #vaccineinjured"; "What are the pros- and cons and possible benefits? Good products and dosage? #MedTwitter #LongCovid #MECFS"; "Stop psychologozing [sic] a physical illness! Help #MedTwitter!!"). The #MedTwitter community also constitutes the common space of a community that shares the same interests or condition. This is why in some cases a pragmatic discourse marker, *hey*, is used to open a conversation addressed to the community funnelled into the hashtag #MedTwitter ("Hey #medtwitter - need actionable advice because…"; "Hey #MedTwitter and esp #cardiology - if you interact with a #LongCovid patient…"; "Hey #MedTwitter #LongCovid representatives are standing by…"; "Hey #MedTwitter, I'm having some chest/heart pain and I suspect it may be…"; "Hey, #MedTwitter & #NeuroTwitter! One of #LongCovid's most serious symptoms is #MentalIllness…"; "Hey #MedTwitter #Cardiotwitter Seeking a few panelists for our upcoming town hall on cardiovascular diseases…"; "Hey #MedTwitter , before this hellsite goes down, does anyone have printable patient education leaflets…").

The hashtag *#NEISvoid* (581 occurrences) is used in similar ways as #MedTwitter. It is designed to gather a community that shares the same set of chronic pathologies. Unlike *#MedTwitter*, which could be intercepted or fostered by specialists in the medical field, *#NEISvoid* was created in March 2020 by a user (whose digital pseudonym is *distracted ghost, @bennessb*) wishing to bring together people with the same condition and feelings. The aim of this hashtag is to gather and index content with the same aim under the same digital marker, as well as to try to unite people suffering from the same condition for support reasons. The occurrences found in this corpus confirm that people experiencing (or talking about) Long Covid

as a chronic issue deserved to "enter" this community by adding the *#NEISvoid* hashtag to their tweets, so that useful information or experiences can be shared even asynchronously.

6. Final remarks

Like the broader discourse on Covid-19, Long Covid is a delicate issue. It shares similar language aspects with Covid-19 but also exhibits rather specific language features. The analysis of the corpus of tweets about Long Covid reveals a variety of polarised attitudes about this condition, as well as different ways of representing the syndrome depending on the intention, addresser and intended audience or network. From a frequency-based standpoint, scientifically based information tends to prevail over negative and critical views. This fact could reflect the nature of Twitter/X, which does not favour the creation of networks based on personal acquaintance, encouraging the development of interest-based circles instead. Users' names are anonymous or hidden under pseudonyms, thus favouring "multiple flattened audiences" (Brandtzaeg and Lüders 2018, 2). At the same time, the use of less specialised, more accessible phrases to popularise medical facts has been a growing trend for some years, increasingly blurring the lines between popular and specialised terminology (see Grego 2023). Private, even intimate experiences are given voice through accounts of personal Long Covid-related experiences. Here, users who share the same pathology tend to describe their health condition as a form of personal narration, but also as a way to find peers in attempting to grow a community with supportive aims.

Twitter/X is a virtual space in which many asynchronous forms of interaction and access are found. Semantic order is achieved by means of device-specific tools such as hashtags and retweets, as we see here in the use of Long Covid-related hashtags. Further research is strongly advised, e.g., focus on the same topic with different timespans; investigation in a diachronic perspective to observe possible variations of language patterns associated with Long Covid; or retrieval processes performed following a debatable piece of news that could spark personal and polarised views (Meledandri 2023). Studies including other variables in the analysis of the corpus (e.g., including retweets to amplify engagement and related forms of discourse) would also be helpful. Furthermore, analysis of infrequent patterns and tokens may reveal new language patterns associated with the topic. Finally, different language layers should be further investigated (lexical, syntactic, device-specific, ideological, etc.) to formulate insights about this discourse.

Francesco Meledandri *holds a PhD in Translation Theory and Praxis: European Intercultural Studies awarded by the University of Bari (2011). After a Research fellowship (2013), he has been a lecturer (English language and translation, 2013-2021) at University of Bari, Dept. of Humanities, along with ICT solutions for Translation (2016-2021), and IT (advanced course, 2018-2021) at the same University. Since January 2022, he has been appointed Research Fellow at Department of Humanities, University of Bari. His research interests focus on both language and translation issues covering different topics, such as the language of sports in both linguistic- and translation-related terms; furthermore, he has been analysing the growing interest towards communication via social networking platforms in order to retrieve real-time, aggregated bulk of data aimed at investigation of micro- and macro-features of language. Other research areas include Translation Studies, CAT Tools, and Machine Translation.*

Works cited

Al-Amer, Rasmieh, et al. "Emotional Reaction to the First Dose of COVID-19 Vaccine: Postvaccination Decline in Anxiety and Stress among Anxious Individuals and Increase among Individuals with Normal Prevaccination Anxiety Levels." *Journal of Personalized Medicine* 12.912 (2022).

Andrews, Nick, et al. "COVID-19 Vaccine Effectiveness against the Omicron (B.1.1.529) Variant." *The New England Journal of Medicine* 386 (2022): 1532-1546.

Nelson, Anitra. "COVID-19: Capitalist and Postcapitalist Perspectives." *Human Geography* 13.3 (2020): 305-309.

Anthony, Lawrence. "AntConc" (Version 3.4.3.0) [Computer Software]. Waseda University, Tokyo.

---. "AntConc (Version 3.5.8)" [Computer Software]. Waseda University, Tokyo.

Awoyemi, Toluwalase, et al. "Twitter Sentiment Analysis of Long COVID Syndrome." *Cureus* 14.6 (2022): e25901.

Azcue, Naiara, et al. "Brain Fog of Post-COVID-19 Condition and Chronic Fatigue Syndrome, Same Medical Disorder?" *Journal of Translational Medicine* 20.569 (2022).

Badaan, Vivienne, et al. "Ideological Asymmetries in Online Hostility, Intimidation, Obscenity, and Prejudice." *Nature Scientific Reports* 13 (2023): 22345.

Baker, Paul. *Using corpora in Discourse Analysis*. London: Continuum, 2006.

Ben-Elia, Eran and Feng Zhen. "ICT, Activity Space–Time and Mobility: New Insights, New Models, New Methodologies." *Transportation* 45 (2018): 267-272.

Birchall, Clare and Peter Knight. *Conspiracy Theories in the time of COVID-19*. London: Routledge, 2022.

Brandtzaeg, Petter B. and Marika Lüders. "Time Collapse in Social Media: Extending the Context Collapse" *Social Media+Society* 4.1 (2018): 1-10.

Brooker, Phillip, Julie Barnett and Timothy Cribbin. "Doing Social Media Analytics." *Big Data & Society* 3.2 (2016): 1-12.

Burgess, Hayley, et al. "COVID-19: The Vaccine Race Continues." *Healthcare Journal of Medicine* 2.2 (2021): 81-91.

Callard, Felicity and Elisa Perego. "How and Why Patients Made Long Covid." *Social Science & Medicine* 268 (2021):113426.

Chaiuk, Tetyana and Olha V. Dunaievska. "Fear Culture in Media: An Examination on Coronavirus Discourse." *Journal of History Culture and Art Research* 9.2 (2020): 184-194.

Cosentino, Gabriele. *Social Media and the Post-Truth World Order: The Global Dynamics of Disinformation*. Cham: Palgrave, 2020.

Cucinotta, Domenico and Maurizio Vanelli. "WHO Declares COVID-19 a Pandemic." *Acta Biomed* 91.1 (2020): 157-160.

D'Ancona, Matthew. *Post-Truth: The New War on Truth and How to Fight Back*. London: Ebury Press, 2017.

Damota, Mulugeta Deribe. "The Effect of Social Media on Society." *New Media and Mass Communication* 78 (2019): 7-11.

Davis, Hannah, et al. "Long COVID: Major Findings, Mechanisms and Recommendations." *Nature Reviews Microbiology* 21 (2023): 133-146.

Dolatabadi, Elham, et al. "Using Social Media to Help Understand Long COVID Patient Reported Health Outcomes: A Natural Language Processing Approach." *J Med Internet Res*. (2023): 25:e45767.

Garfin, Dana Rose, Roxane Silver Cohen and Alison E. Holman. "The Novel Coronavirus (COVID-2019) Outbreak: Amplification of Public Health Consequences by Media Exposure." *Health Psychology* 39.5 (2020): 355-357.

Georgalou, Mariza. *Discourse and Identity on Facebook*. London: Bloomsbury, 2017.

Grego, Kim. "COVID-19 'Brain Fog' Narratives in Institutional, News Media and Academic Sources." *Communicating medical science in the digital age: culture, knowledge, expertise, practices, 2nd International CIRLaM Conference 25-26-27 May 2023, Naples, Italy*. Conference paper.

Hinton, Sam and Larissa Hjorth. *Understanding Social Media*. London: Sage, 2013.

Hoffmann, Christian R. and Wolfram Bublitz, edited by. *Pragmatics of Social Media*. Berlin: De Gruyter, 2017.

Kapoor, Kawaljeet Kaur, et al. "Advances in Social Media Research: Past, Present and Future." *Information System Frontiers* 20 (2018): 531-558.

Katella, Kathy. "Comparing the COVID-19 Vaccines: How Are They Different?" *Yale Medicine* 24 April 2024. https://www.yalemedicine.org/news/covid-19-vaccine-comparison.

Khan, Madiha. "#MedTwitter: The Good, the Bad, and the Surprisingly Useful." *Nejm Journal Watch* 5 April 2022. https://blogs.jwatch.org/general-medicine/index.php/2022/04/medtwitter-the-good-the-bad-and-the-surprisingly-useful/.

KhosraviNik, Maijd. "Digital Meaning-Making across Content and Practice in Social Media Critical Discourse Studies." *Critical Discourse Studies* 19.2 (2022): 119-123.

Knoll, Maria Deloria and Chizoba Wonodi. "Oxford–AstraZeneca COVID-19 Vaccine Efficacy." *The Lancet* 397.10269 (2021): 72-74.

Li, Junjiang and Philippe Giabbanelli. "Returning to a Normal Life via COVID-19 Vaccines in the United States: A Large-scale Agent-Based Simulation Study." *JMIR Med Inform* 9.4 (2021): e27419.

Liu, Sixiao and Haoran Chu. "Examining the Direct and Indirect Effects of Trust in Motivating Covid-19 Vaccine Uptake." *Patient Education and Counseling* 105 (2022): 2096-2102.

Mayo Clinic. "Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS)." https://www.mayoclinic.org/diseases-conditions/chronic-fatigue-syndrome/symptoms-causes/syc-20360490.

McIntyre, Lee. *Post-Truth*. Cambridge: MIT Press, 2018.

Meledandri, Francesco. "A Jab to Medical Facts: Sentiment of the Twittersphere in COVID-19 Vaccination Official Discourse: A UK-Italy Comparison." *ESP across Cultures* 20 (2024, in press).

---. "Out-of-the-(ballot)box: Legitimation of a New Popular Will in Brexit-related Social Media Engagement." *Thinking Out of the Box in English Linguistics, Language Teaching, Translation and Terminology. Proceedings of the XXIX AIA Conference*. Edited by Katherine Ackerley, et al. Padova: Padova University Press, 2023. 157-182.

Pedrini, Giulia. "'The 'Vaccine Race': Metaphorical Conceptualizations of the Search of an Immunization Against COVID-19." *International Journal of Translation* 33 (2021): 129-152.

Polack, Fernando, et al. "Safety and Efficacy of the BNT162b2 mRNA COVID-19 Vaccine." *The New England Journal of Medicine* 383.27 (2020): 2603-2615.

Saladino, Valeria, Davide Algeri and Vincenzo Auriemma. "The Psychological and Social Impact of COVID-19: New Perspectives of Well-Being." *Frontiers in Psychology* 11 (2020): 1-6.

Schulman, Roy. "COVID-19 and the Post-Truth Age: The Role of Facts in Public Policy" *Online conference,* 15 June 2020: https://www.inss.org.il/publication/coronavirus-and-post-truth-seminar/.

Shaw, David. "Invisible Enemies: Coronavirus and Other Hidden Threats." *Journal of Bioethical Inquiry* 17 (2020): 531-534.

Stefanowitsch, Anatol. *Corpus Linguistics: A Guide to the Methodology*. Berlin: Language Science Press, 2020.

Sussman, Kristen, et al. "COVID-19 Topics and Emotional Frames in Vaccine Hesitation: A Social Media Text and Sentiment Analysis." *Digital Health* 9 (2023): 1-11.

Tamássy, Réka and Zsuzsanna Géring. "Rich Variety of DA Approaches Applied in Social Media Research: A Systematic Scoping Review." *Discourse & Communication* 16.1 (2022): 93-109.

United Nations. "United Nations Comprehensive Response To COVID-19. Saving Lives, Protecting Societies, Recovering Better. 2021 update." https://reliefweb.int/report/world/united-nations-comprehensive-response-covid-19-saving-lives-protecting-societies.

US Centers for Disease Control and Prevention. "ME/CFS Basics." https://www.cdc.gov/me-cfs/about/index.html.

Voysey, Merryn, et al. "Safety and Efficacy of the ChAdOx1 nCoV-19 Vaccine (AZD1222) against SARS-CoV-2: An Interim Analysis of Four Randomised Controlled Trials in Brazil, South Africa, and the UK." *The Lancet* 397.10269 (2021): 99-111.

Wong, Timothy and Danielle J. Weitzer. "Long COVID and Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS)—A Systemic Review and Comparison of Clinical Presentation and Symptomatology." *Medicina* 57:418 (2021).

World Health Organization. "A Clinical Case Definition of Post Covid-19 Condition by A Delphi Consensus, 6 October 2021" *World Health Organization*. https://www.who.int/publications/i/item/WHO-2019-nCoV-Post_COVID-19_condition-Clinical_case_definition-2021.1.

---. "Post COVID-19 Condition (Long COVID)." https://www.who.int/europe/news-room/fact-sheets/item/post-covid-19-condition.

---. "Statement on the Fifteenth meeting of the IHR (2005) Emergency Committee on the COVID-19 Pandemic." https://www.who.int/news/item/05-05-2023-statement-on-the-

fifteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic.

Zappavigna, Michele. "Ambient Affiliation: A Linguistic Perspective on Twitter." *New Media Society* 13 (2011): 788-806.

---. *Discourse of Twitter and Social Media: How We Use Language to Create Affiliation on the Web*. London: Continuum, 2012.

---. "Searchable Talk: The Linguistic Functions of Hashtags." *Social Semiotics* 25.3 (2015): 274-291.