A Chinese Virus? Uncovering The Linguistic and Political Implications of Naming a Disease Through a Corpus-Study of Covid-19

Luz Andrea Alvariño Gabeiras

Albert-Ludwigs Universität Freiburg

Distant Reading

Prof. Dr. Christian Mair

September 22, 2020

Table of Contents

1. Introduction

In an ill-fated act of foreshadowing, Michael Rosen began his 2018 podcast 'Naming Diseases' by asking "anyone got the flu and is it Japanese or Aussie flu?" and stating that "when a new disease emerges, one of the first tasks will be deciding what to call it" (Rosen et al.). As it turns out, he was right, the new disease has emerged and this time, it's a Chinese flu. As we have come to find out, the name given to a particular disease can be just as infectious and propagate just as quickly, and sometimes, be nearly as harmful than the disease itself. Although most disease names we are familiar with have become opaque with time, the emergence of a new disease presents an unique opportunity for observing the processes associated with disease naming, and the changes that a disease name undergoes in its early stages. Most people unaffiliated with the medical field would not be able to tell what the name 'Naples Soldier' originally meant, but the name was once not only extremely transparent, but a deliberate choice backed by powerful political interests. Since the beginning of the ongoing Covid-19 pandemic, we have seen names for the disease arise and be shot down by international backlash, or defended by scientific organizations. The present diachronic study will seek to make sense of this onomastic mess, exploring first whether terms for Covid-19 should even be considered proper names, and then looking at the evolution of naming for Covid-19, from SARS-CoV-2 to Kung-Flu.

The study will begin by presenting the theoretical concepts surrounding the naming of diseases, and particularly Covid-19, as well as the study of proper names in Chapter One; a summary of the data collected and the research methods used to examine it in Chapter Two; an overview of the analysis and the results obtained in Chapter Three, and an explanation and summary of such results in the Chapters Five and Six.

2. Literature Review

This chapter aims to give a theoretical overview of the process of disease naming, starting with a brief historical introduction followed by a summary of the names that have been given to Covid-19 and how they came to be; and finally, offering an account of names as the object of study of onomastics within linguistics, including corpus and socio-onomastics.

2.1. Naming a Disease

History of Disease Names

In order to understand how the names for Covid-19 have come to be and what their possible effects could be, one must first get an understanding of how the process of naming a disease has changed throughout history as well as the recent attempts at regulating such processes. "Names are a linguistic universal … all known languages make use of names—most commonly, but not exclusively, to identify individual people and places" (Hough 2016: 1). Names of diseases are no different, seeing as "when a new threat to life emerges, the first and most pressing concern is to name it" (Organization 2020c). In the case of infectious diseases in particular, giving them a name as fast as possible allows them to go from a mysterious, invisible danger to a well-defined entity, which is usually much less "terrifying" (Rosen et al. 2018). Although this urgency in naming diseases has persisted in time, the process of naming diseases has changed considerably. In fact, looking at disease names can serve as a window into the anxieties, customs and interests of a particular age.

Seven main ways of naming diseases can be recognized through history, diseases were named after: (1) the doctor who first discovers or describes the disease; (2) the first patient diagnosed with the disease; (3) the place where it originates; (4) the group affected; (5) the perceived source of the virus (e.g. pigs); (6) the agent originating the disease (e.g. streptococcus); (7) the perceived cure of the disease; (7) the symptoms of the disease; and

sometimes (8) fictional characters or elements associated with the disease (Sykes 1896; Majewska 2014; Rosen et al. 2018).

The method chosen might depend on who names the disease. Traditionally, names were suggested primarily by physicians, writers and philosophers, who "[forged] new names as blacksmiths forge horseshoes" (Sykes 1896: 1009), as in the case of Scarlatina, named by Doctor Thomas Sydenham Sydenham in his work "*Observationes Medicarum*" (Sykes 1896:1007). Although the names came primarily from medical professionals, the political interests of the country where they originated were often the ruling factor over the decision. Such is the case of the Spanish Flu, named after Spain merely because France, Britain, and America – where the disease probably originated – were at war, and censored their newspapers to keep their soldiers' morale high. An even better example is the naming of Syphilis, which was called the 'Italian Disease' in France, the 'French Disease' in Italy, the 'Spanish Disease' in the Netherlands, the 'Christian Disease' in Turkey and the 'Polish Disease' in Russia, depending on which nations were warring with which (Rosen et. al. 2018). In other cases, the motivation came from the physician's own career,

> at the beginning of the 20th century, it was considered a great honor to have a microbe, a deadly microbe named after you. Hence, we have listeria after the surgeon Lyster. (Wright in Rosen et al.)

This tendency towards scapegoating and self-laureation is what Spinney (2017) calls "the ugly history of disease naming." Unfortunately, its effects have dragged on well into modern times, with AIDS being originally named 'Gay Related Immune Deficiency' (GRID), or 4H, standing for Haitian Homosexuals, Hemophiliacs and Heroin Addicts (Rosen et.al. 2018). Moreover, the flu pandemic of 2009 was then called 'swine flu' in spite of being propagated by humans instead of pigs, which resulted in a government-mandated killing of all of Egypt's pigs (Spinney 2017).

Recently, new attempts by institutions such as the World Health Organization (WHO) have been made at regulating these historical shortcomings of disease naming in order to avoid discrimination and harm. These attempts have supposed the biggest motor of change for the present conventions of naming ailments. The first was the "taxonomic scheme" created by 'The Coronavirus Study Group (CSG)' of the 'the International Committee on Taxonomy of Viruses (ICTV)' in 1966 (Lesney 2020), which states that

> The universal virus classification system shall employ the hierarchical levels of realm, subrealm, kingdom, subkingdom, phylum, subphylum, class, subclass, order, suborder, family, subfamily, genus, subgenus and species. (Lesney 2020)

The *World Health Organization Best Practices for the Naming of New Human Infectious Diseases* was introduced in May of 2015 and outlines the standards seen in Figure 1.

Figure 1

Guidelines for Disease Naming Introduced by WHO.

Disease names may include:	Examples of useful terms
Generic descriptive terms (clinical symptoms, physiological processes, and anatomical or pathological references/systems affected)	• Respiratory, neurologic, hemorrhagic • Hepatitis, encephalitis, encephalopathy, diarrhoea, enteritis, immunodeficiency, palsy • Pulmonary, cardiac, gastrointestinal, spongiform • Syndrome, disease, fever, failure, deficiency, insufficiency, infection
Specific descriptive terms:	
Age group, population of patients	juvenile, pediatric, senile, maternal
Time course, epidemiology, origin	Acute, sub-acute, chronic, progressive, transient, contagious, congenital, zoonotic
Severity	Severe[A], mild
Seasonality	Winter, summer, seasonal
Environment	Subterranean, desert, ocean, coastal, river, swamp
Causal pathogen and associated descriptors	• Coronavirus, salmonella/salmonellosis, influenza virus, parasitic • Novel[B], variant, reassortant • Subtype, serotype
Year (+/- month) of first detection or reporting[C]	2014, 3/2014
Arbitrary identifier	Alpha, beta, a, b, I,II,III, 1,2,3

Disease names may NOT include:	Examples to be avoided
Geographic locations: Cities, countries, regions, continents	Middle East Respiratory Syndrome, Spanish Flu, Rift Valley fever, Lyme disease, Crimean Congo hemorrhagic fever, Japanese encephalitis
People's names	Creutzfeldt-Jakob disease, Chagas disease
Species/class of animal or food	Swine flu, bird flu, monkey pox, equine encephalitis, paralytic shellfish poisoning
Cultural, population, industry or occupational references	Occupational, legionnaires, miners, butchers, cooks, nurses
Terms that incite undue fear	Unknown, death, fatal, epidemic

Note: Reprinted from *WHO best practices for naming of new human infectious diseases* (p.2) by World Health Organization, 2015.

These conventions suggest highly scientific names for diseases, including medical names for the symptoms as well as numbers and letters when necessary, and avoiding personal or geographical names. In the past, names of this sort suggested by scientists were met by skepticism or simply failed to meet the necessary 'catchiness' requirements, succumbing to the more 'folk' names suggested by the common population, as in the case of "la grippe," a popular word in France for the flu that substituted the more scientific British "influenza" (Sykes 1896: 1008). Rosen et.al. describe a similar phenomenon in the present, stating that carefully orchestrated, scientific names are often incomprehensible for the average person, "cannot be tweeted", almost invariably become opaque acronyms and sometimes even

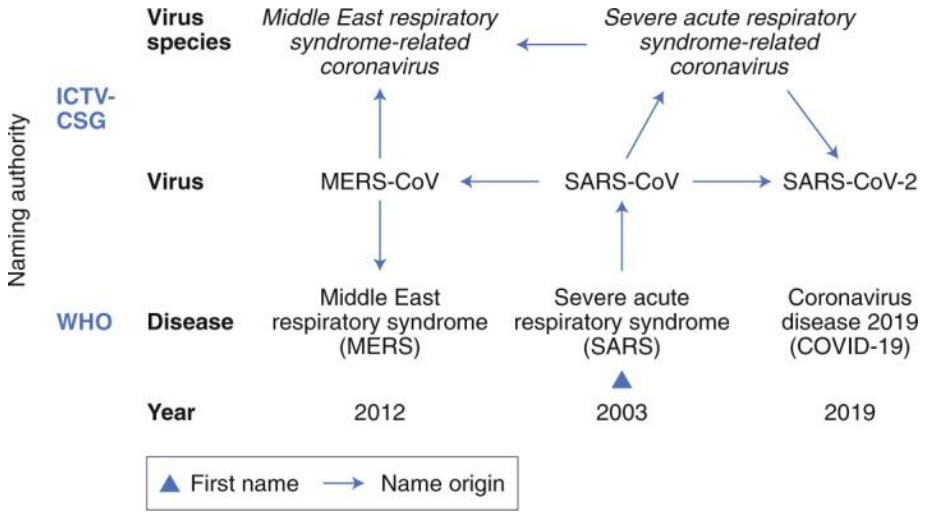
obscure important information necessary to avoid the disease. Therefore, question remains as to whether the indications provided by WHO will be helpful in reducing the potential harm and discrimination coming from disease naming or if the guidelines will be ignored by the bulk of the population in favor of more easily remembered, and 'tweetable' names. Nevertheless, these conventions have undoubtedly modified the most recent diseases discovered, such as Covid-19, as we will see in the following section.

Names for Covid-19

During the month of December of 2019, Dr. Zhang Jixian of the Hubei Provincial Hospital of Integrated Chinese and Western Medicine in Wuhan, started seeing suspicious cases of pneumonia which affected multiple members of a single family, pointing at an infectious disease. After alerting authorities, it was discovered that the cause was what was then named a 'novel coronavirus' (Wei, Cheng & Meng 2020). On the 9th of January of 2020, WHO put out their first statement on the new pneumonia outbreak, also naming it 'novel (or new) coronavirus' (Organization 2020d). This was the first scientific name for Covid-19, and by January 20th, it had already been abbreviated to '2019-nCoV' in the first of what would eventually be a total of 209 situation reports by WHO (Organization 2020a). The name would eventually be changed to '2019-nCoV acute respiratory disease' in compliance with the 2015 WHO guidelines (Organization 2020b). The virus was subsequently called 'Human Coronavirus 2019 (HCoV-19)' (Wong et al. 2020) and finally, after being sequenced and classified, it was given its current formal names 'SARS-CoV-2' in accordance with the CSG guidelines, and Covid-19 in accordance with the WHO guidelines, as seen in Figure 2 (Gorbalenya et al. 2020).

Figure 2

Naming of Coronavirus Diseases Through History.



Note: Reprinted from *The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2* (p.537) by the Coronaviridae Study Group of the International Committee on Taxonomy of Viruses (2020).

However, by the time these scientific terms had come to be, popular terms had already emerged in order to name the disease. The first of these referenced the geographical origin of the virus, such as 'China virus', 'Chinese virus' or 'Wuhan virus' (Hui 2020); the magazine *Nature*, for instance, first used the term 'China coronavirus' (Callaway & Cyranoski 2020) in their initial report of the disease and later apologized, citing the WHO guidelines and admitting that geographically-based names might contribute to fear and discrimination (Shu 2020). Additionally, abbreviations of the names 'Coronavirus' and 'Covid-19' started appearing, with the popular 'Corona', 'Covid' and most colloquial of all, 'Rona' or even 'Miss Rona' (Shariatmadari 2020). The virus also attracted more controversial names, echoing the ages long tradition of using disease names as a tool for political scapegoating, like 'Kung-Flu' (Jiang 2020); or 'Communist Virus', used by figures like journalist Kevin Libin,

who suggested that "perhaps we can all compromise and call it the 'Chinese communist virus.' Chicom-19?" among others (Libin 2020; Henninger 2020; Jordan 2020).

In summary, the following names have been predominantly used for Covid-19[1]:

Table 1

Names for Covid-19

Type	**Names**
Scientific	Coronavirus, Novel coronavirus, 2019-n-CoV, SARS-CoV-2, Covid-19
Abbreviations	Corona, Covid
Colloquial	Rona, Miss Rona
Geographically based	China virus, Chinese virus, Wuhan virus, Kung-Flu
Politics based	Communist virus

[1] The most salient of these names will serve as the main focus of the analysis later conducted in chapters 3 to 6.

2.2. The Study of Names

Defining Names

Within linguistics, onomastics is the discipline concerned with the study of names and naming. Although the study of proper names has traditionally focused on personal and place names, the field covers a wide expanse of topics, including the naming of conflicts, disasters and diseases. Within onomastics, names can be defined in various ways and authors like Searle (1958)[2], Long (1969), Donnelan (1972), Burne (1973), or Bredart et.al. (1996) have extensively discussed what can and cannot be considered a proper name, as well as the various ways of classifying proper names according to their syntactic, semantic and pragmatic properties. In *Names and Grammar* (2016) Langendonck and Velde state that "names are definite nouns with unique denotation that display an inherent basic level sense" (Langendonck & Velde 2016: 35):

- 'Definite' alluding to the fact that names refer to an existing entity which is unique. The definiteness of names can be syntactically 'diagnosed', for instance, if a noun can appear, at least colloquially, in right dislocation as in: "he's an idiot, *John*". (Langendonck & Velde 2016: 40–41; emphasis added)
- Names' unique denotation refers partly to their definiteness, but also to their "incompatibility with restrictive relative clauses and their inability to refer back anaphorically" (40); for example, 'restrictive modifiers' such as "that" cannot be used with names, as in the case "*Ghent* **that** is the most beautiful city in Flanders, was one of the biggest in medieval Europe" as opposed to "*the city* **that** I visited was nice" (Langendonck & Velde 2016: 41–42; emphasis added).
- The inherent basic sense of names refers to their "inherent categorical presupposed sense," for instance in the sentence "*London* is on the Thames," it is presupposed that

[2] Later summarized by Evans & Wimmer (1990).

> London is a city at the most basic level. In the case of names for Covid-19, the prototypical basic sense would be 'disease' or 'virus' (Langendonck & Velde 2016: 42–43; emphasis added).

To these basic tenets, a few more characteristics can be added. Concerning orthography, proper names are usually "marked with initial capitalization (oxford-Oxford)." (Motschenbacher 2020: 90) Morphologically, "a proper name prototypically consists of a singular proper noun that does not allow for pluralization (Oxford – Oxfords)" although "less prototypical proper names that are collective plural forms do not allow for singularization (the Alps – the Alp)." (90) And lastly, syntactically, the absence of a premodifying article (or any other determiner) is typical of proper names (Oxford – the Oxford)"; and proper names cannot "express contrastive definiteness (Oxford – an Oxford)" (90).

Onomastic Classification of Names

In their abridged typology (2016), Langendonck & Velde classify nouns with the aforementioned characteristics in seven categories: (1) 'Personal Names', for prototypical names of people; (2) 'Place-names' for names of cities, countries, regions and other geographical phenomena; (3) 'Names of Months'; (4) 'Trade and Brand Names' for names of companies; (5) 'Numbers', a more controversial category for cases in which numbers act as nouns, such as "*three* is a sacred number"; (6) 'Names of Diseases and Biological Species', the most relevant for the present study, including names of ailments such as "*Aids*" or "*Ebola*"; and finally (7) 'Autonyms' for those cases in which "linguistic expressions refer to themselves" as in "*bank* is a homonymous word". (Langendonck & Velde 2016: 50–54; emphasis added)

Within these, disease names is among the least prototypical categories[3], as some of its members behave like a common noun rather than a proper name (Langendonck & Velde 2016: 50), implying that not all of the terms used to designate a disease may be considered disease names. Langendonck & Velde state that "diseases are apparently never Names in English" (Langendonck & Velde 2016: 53), and only sometimes in languages like Dutch. The authors cite familiarity as a marker for 'nameness' and capitalization as the main way to differentiate disease names from common nouns (Langendonck & Velde 2016: 53):

- "The *Aids* disease expands in Africa." (Langendonck & Velde 2016: 53)
  vs.
- "The *influenza* disease returns every year." (54)

In contrast, Abel (2014a, 2014b, 2018) implies that eponyms and toponyms may be disease names in English, and that words that formerly were eponyms, such as 'syphilis'[4] can also retain some name-like characteristics, such as the difference in capitalization.

How the dichotomy between common nouns and proper names in disease names concerns terms for Covid-19 is, perhaps, even trickier, due to the high number of words being used to name disease, and their low grade of familiarity in the English language. An attempt to clarify whether they behave or not like proper names will be made later in the present paper.

[3] It has been suggested that "names have prototype effects" meaning "particular (groups of) names may show more or fewer of the grammatical and semantic features generally considered to be typical of names" (Motschenbacher 2020: 90).

[4] "from Syphilis, sive Morbus Gallicus, the title of a Latin poem (1530), from the name of the character Syphilus, the supposed first sufferer of the disease." ("syphilis" 2019)

Corpus-Onomastics and Socio-Onomastics

As it should become apparent by the previous section, the characteristics and classification of proper names remains a point of contention, particularly when differentiating them from common nouns and especially in less prototypical categories such as disease names. Motschenbacher (2020) signals at the lack of corpus linguistic studies as a possible reason for this theoretical battleground, pointing out that

> we still have a limited picture of how proper nouns are actually used, but also that the descriptions we have are strongly influenced by notions of how proper nouns should be used in accordance with standard language norms (Motschenbacher 2020: 94);

Corpus studies could present an opportunity to study how proper nouns are used in everyday discourse, as well as to "empirically verify the linguistic characteristics of names as they surface in actual language use, especially at the semantic and grammatical levels" (Motschenbacher 2020: 89). As examples of corpus-onomastics studies, Motschenbacher (2020: 96-97) points out Tse's (2002; 2003; 2004; 2005) works, most of which deal with the placement of a definite article before proper names.

Within onomastics, socio-onomastics focuses on the point of relation between names and society, and "can be defined as a sociolinguistic study of names," exploring the "use and variation of names" in relation to the "social, cultural, and situational field in which names are used" (Ainiala 2016: 413). Its main focus is synchronic and diachronic name variation (415), although most studies conducted so far, such as the ones compiled in Ainiala & Östman (2017) have primarily focused on 'Personal Names' or 'Place-Names', with hardly any research going into the study of 'Disease Names'[5].

[5] With the exception of Majewska (2014).

3. Data and Methodology

This chapter gives a survey of the data collected in the course of the study, as well as the methods used in the present study. Corpus methods will be applied to the study of names used for Covid-19, alongside a study of the social circumstances in which the names are used; therefore, the framework and methods of corpus-onomastics and socio-onomastics will be applied throughout the analysis.

3.1. Data Collection Process

The first step of the analysis was to collect data from Mark Davies' Coronavirus Corpus. The corpus, currently containing 601 million words, exists as a subset of the NOW Corpus, which compiles "data from web-based newspapers and magazines from 2010 to the present time" (Davies 2016), this subset includes:

> 1. Articles that have at least two occurrences of the words {coronavirus, COVID, or COVID-19}.
>
> 2. Articles that have one of the following words/strings in the title: at-risk, cases, confirmed, contagious, containm*, coronavirus, covid*, curbside, curve, deaths, disinfect*, distanc*, epicenter, epidemic, epidemiol*, flatten*, flu, high-risk, hoard*, hospital*, hydroxychloroquine, infect*, influenza, isolat*, lockdown, lock-down, mask*, nursing, outbreak, pandemic, panic, patient*, pneumon*, preventative, preventive, quarantin*, re-open*, reopen*, respiratory, sanitiz*, self-isolat*, shelter*, shutdown, spread, spreading, stay-at-home, stay at home, stockpil*, testing, vaccine*, ventilator*, virus. (Davies 2020)

In addition, the corpus contains data starting from January 2020, and is updated every ten days.

Twelve of the most salient names for Covid-19 were selected for the analysis: *Coronavirus, SARS-CoV-2, Covid-19, Corona, Covid, Rona, Miss Rona, China virus, Chinese virus, Wuhan virus, Kung-Flu, and Communist virus.* First, a random sample of twenty concordance lines for every name was obtained. Then, using the chart function, it was possible to collect a total normalized frequency (per million words) of each of the names at 10-day intervals starting from January 1st until September 20th. This process rendered a total of 289 observations, recorded in a data collection table (Table 2).

Table 2

Data Collection Table Sample

Name	Type[6]	Date (y/m/d)	Frequency (PMW)
Coronavirus	Scientific	20-01-01	2,733.29
Coronavirus	Scientific	20-02-01	4,943.78

Which was subsequently processed, both manually and using RStudio (RStudio Team 2016).

3.2. Research Methods

The initial objective of this paper was to observe the evolution of the different names given to Covid-19 in the course of the pandemic. After a review of the previous literature both on the naming of diseases and the study of proper names, the following two research questions were outlined:

I. As introduced in Section 2.2 the category of disease names is controversial within Onomastics, due to many of its members behaving in non-prototypical ways. The first question extracted from this principle is whether any names used for Covid-19 can be considered proper names (fulfilling the characteristics outlined in section 2.2), or rather, if they behave as common nouns.

[6] See Table 1.

II. Whether the popularity of names for Covid-19 has evolved in favor of scientifically-suggested names such as SARS-CoV-2; geographically-based names that, according to WHO, invoke racist connotations; or catchy and popular abbreviations – and how this evolution in popularity relates to the socio-political circumstances of the time.

In order to address these questions, the analysis was structured as follows:

I. A qualitative analysis of twenty concordance lines each for the names ‘Covid-19’ and ‘Wuhan virus’, respectively, was conducted. The two names were picked due to the disparity between them; where Covid-19 was suggested by the scientific community and aims to be as neutral and semantically sterile as possible, Wuhan virus is a popular name that directly alludes to the geographical origin of the virus, but both are headed by nouns, unlike for instance ‘Chinese virus’.

II. A quantitative analysis of the selected names for Covid-19, focused on total normalized frequency counts for each of the 10-day intervals collected. In order to contextualize key points in the variation, a few examples were extracted from the concordance lines.

The resulting analysis will be presented in the succeeding chapter, and its results broken-down and summarized in chapters five to six.

4. Analysis and Results

This chapter presents a review of the results, beginning with the qualitative evaluation of concordance lines for ‘Covid-19’ and ‘Wuhan virus’ and concluding with the quantitative analysis of the evolution of names for Covid-19.

4.1. Covid-19 vs. Wuhan Virus: Name or Noun

Starting with Covid-19, the following random sample of twenty concordance lines was obtained:

Figure 3

Concordance lines for Covid-19.

20-03-04 ZA	Covid-19 was officially declared a pandemic by the Director General of the WHO on 11th March
20-03-09 CA	A security guard exits the new COVID-19 clinic at the site of the former Hotel Dieu hospital on March 9, 2020
20-02-23 US	Samsung isn't the first company to shut down factories due to the COVID-19 outbreak
20-03-17 US	now is making sure that we have all the tools to be supportive during this COVID-19 crisis
20-03-16 US	She'd come this far without being infected with COVID-19
20-03-09 --	Over 100 countries have now reported laboratory-confirmed cases of COVID-19
20-03-15 --	remains in Health Protection Condition (HPCON) Alpha in response to the spread of COVID-19 in the state of Colorado
20-03-01 IE	In response, a Covid-19 unit has been set up which is being advised by medical experts
20-03-05 GB	International cooperation is essential to deal with the health and economic impact of the COVID-19 virus
20-02-24 NG	Infection Prevention and Control (IPC)' as part of preparations against COVID-19 in Nigeria and Africa at large
20-03-14 PH	Paranaque LGU beefs up measures against COVID-19 # QUEZON CITY, March 13 (PIA) -- In an effort to contain
20-03-09 PK	percentage of recovered persons is more than 50 per cent that breaks the myth that Covid-19 would kill every infected person
20-03-09 US	That lobbyist had tested positive for COVID-19 coronavirus
20-03-06 US	This week's theme: diseases that wreaked more havoc than COVID-19 currently has
20-03-07 --	to ensure our state has all the necessary resources engaged as we respond to COVID-19
20-03-17 SG	to Keep You Company During Covid-19 Lockdown
20-03-03 US	The FBI and local law enforcement are warning people to be wary of COVID-19 scams
20-02-18 SG	wearing protective facemasks amid fears over the spread of the COVID-19 coronavirus
20-03-04 PH	The Fed's San Francisco regional bank reported that the COVID-19 outbreak has led to decreased demand for aircraft from China
20-03-03 GB	"Covid-19 has the potential to spread extensively"

Little can be said about the orthography and morphology of Covid-19. Orthographically, the term was invariably capitalized in all examples, either only the initial letter as ‘Covid-19’, or all in upper case as ‘COVID-19. It is worth noting, however, that the term never appears in quotation marks. Morphologically, Covid-19 does not seem to allow for pluralization, as would be expected of a proper name.

Its syntaxis, however, is interesting. It appears that the term is never preceded by a determiner is used as a stand-alone as in:

- Example 1: *Covid-19* was officially declared a pandemic.[7]
- Example 2: *Covid-19* has the potential to spread extensively.
- Example 3: Over 100 countries have now reported laboratory-confirmed cases of *COVID-19*.

Only being anteceded by a determiner when part of a noun phrase:

- Example 4: Samsung isn't the first company to shut down factories due to the *COVID-19 outbreak*.

As for more abstract characteristics, Covid-19 seems to be a 'definite' noun, being able to take colloquial right-dislocation (Langendonck & Velde 2016: 40–41 as in the fictional example: 'It's a dangerous thing, Covid-19'. The name refers to a single well-defined entity and has or has an unique denotation as well as the inherent basic sense of 'disease' (Langendonck & Velde 2016: 42–43), although it is not possible to prove syntactically from the existing examples.

The following examples were collected for Wuhan virus:

[7] All examples extracted from the concordance line samples in Figures 3 and 4. Source: Coronavirus Corpus (Davies 2020).

Figure 4

Concordance lines for Wuhan virus.

20-01-31 SG	the coronaviruses, including the Wuhan virus, are mostly found in bats and other warm-blooded animals
20-07-29 US	"Apparently, I have the Wuhan virus," he said
20-01-23 US	Sometimes referred to as the Wuhan virus, it has been temporarily named the "2019-nCoV"
20-02-01 SG	queues at retailers and pharmacies across Singapore, with people buying face masks as the Wuhan virus continues to spread globally
20-03-24 IN	refrained from referring to the virus as the "Chinese virus " or the "Wuhan virus"
20-03-14 SG	Mike Pompeo, referred to it as the "Chinese coronavirus" or the "Wuhan virus"
20-01-29 NG	gold rally will rely on developments for good, or for ill, of the Wuhan virus situation
20-04-13 IN	early stages of the outbreak in January, Chinese state media referred to the "Wuhan Virus" but since then, the use of geographic terms has become highly politicised
20-03-26 IN	In recent weeks, Pompeo has stepped up his use of " Wuhan virus," accusing China of putting the world at risk by not revealing more
20-04-14 NZ	things that are fundamentally wrong with the system that was already breaking down before the Wuhan virus
20-01-28 ZA	Safe haven and commodities react to Wuhan virus risk
20-01-31 NZ	issues behind the Wuhan Virus
20-03-26 US	have taken to referring to the coronavirus as the "China virus" or "Wuhan virus" to highlight China's role in the pandemic
20-05-04 NZ	Pompeo, it may well seem very brave and cutting-edge to use the expression "Wuhan virus" or to call for bigger and bolder rhetorical attacks on China
20-06-11 PK	"Vague concerns about Wuhan virus are still weighing on U.S. stocks,"
20-03-24 --	I have no problem with people calling it the China virus or wuhan virus or covid 19 or sars-co-2 or whatever., just so long as people know
20-07-25 BD	Trump calls the coronavirus "the Wuhan Virus," a reference to the name of the Chinese city where it first
20-02-05 SG	The Wuhan virus spreads much like the common cold, through close contact
20-02-06 AU	US stocks on Monday dropped the most since October as fear surrounding the Wuhan virus prompted rotation from risk assets to safe havens
20-03-16 US	Ruth Bader Ginsburg needs to be very worried about the Wuhan Virus - in addition to whatever ongoing treatment she may be having for cancer

Orthographically, the name Wuhan virus shows more variation Covid-19, it appears initially capitalized, as:

- Example 5: *The Wuhan virus* spreads much like the common cold.

With both words capitalized:

- Example 6: Trump calls the coronavirus *"the Wuhan Virus"*.

As well as with neither:

- Example 7: I have no problem with people calling it the China virus or *wuhan virus.*

Furthermore, when not part of a NP, the term often appears in quotation marks, although not always:

- Example 8: Issues behind the *Wuhan Virus.*
- Example 9: Sometimes referred to as the *Wuhan virus.*

Morphologically, there are no instances of pluralization.

Syntactically, the name is almost always preceded by the definite article 'the', except in:

- Example 10: Vague concerns about *Wuhan virus* are still weighing on US stocks.

An interesting example is:

- Example 11: Pompeo has stepped up his use of "*Wuhan virus*".

In which the word does not actually refer to the virus, but to the name itself. Still, the name never appears with the article 'a' (i.e. A wuhan virus), denoting that it is a definite entity and does not allow for 'contrastive definiteness'[8].

The characteristics outlined by Langendonck & Velde (2016) are harder to determine in the case of Wuhan virus. The fictional example of right-dislocation 'it's a nasty disease, Wuhan virus' seems less intuitive than for Covid-19, but still possible. Like Covid-19, the name's unique sense (i.e. 'disease') and unique denotation ('Covid-19') are quite clear from the examples, even without applying syntactic 'diagnosis' mechanisms.

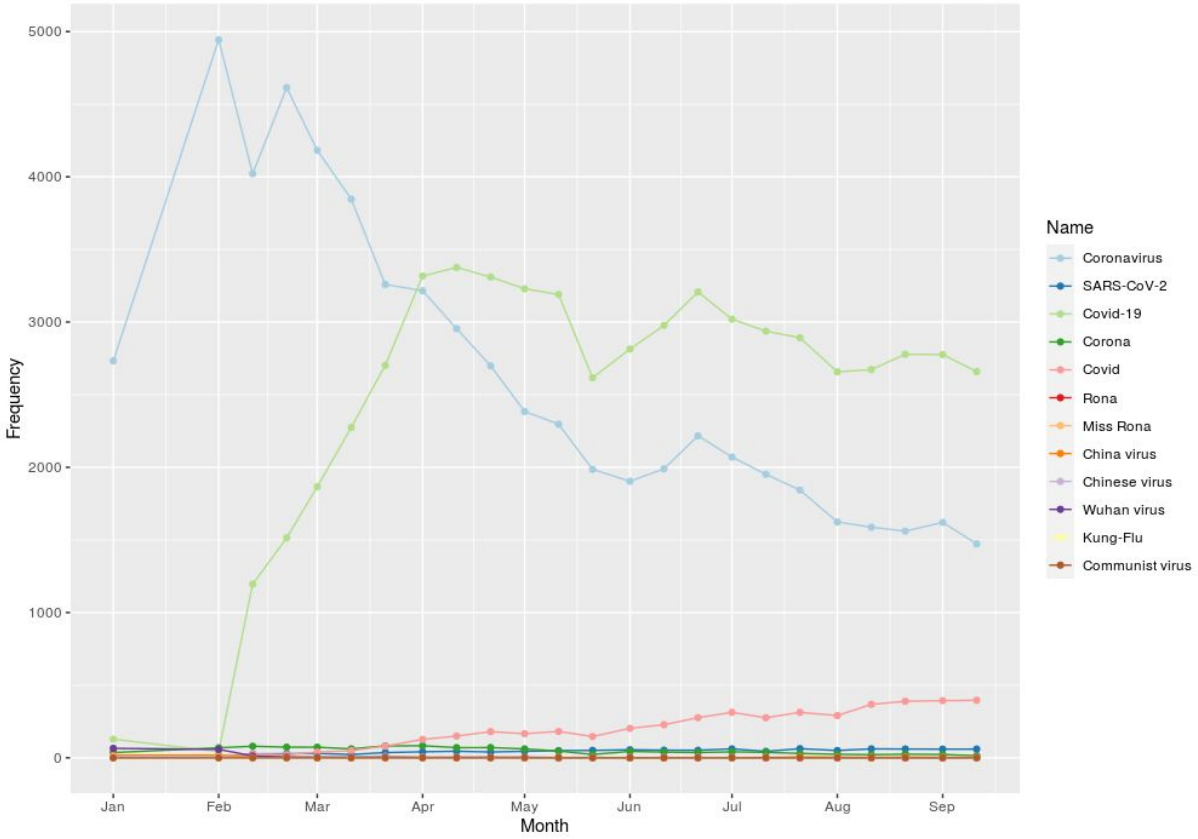
4.2. Variation in the naming of Covid-19

To start with, Figure 5 shows a full breakdown of the normalized frequencies for all twelve names for Covid-19 selected.

[8] Motschenbacher (2020: 90).

Figure 5

Normalized frequency (PMW) of names for Covid-19



The first noticeable trend is the disparity in normalized frequency between two of the names (i.e. Coronavirus and Covid-19) and the rest. Coronavirus started out with the highest normalized frequency in January and then plummeted as the more scientific name Covid-19 – which started out as one of the lowest – took over from April to September. From April onwards an uptick in the use of the abbreviation Covid is visible.

Therefore, April supposed a turning point in the usage of the most popular names for Covid-19; when looking at some concordance lines from this month we find that Coronavirus is starting to be considered a more general term and is often followed by COVID-19:

- Example 12: A Palestinian man wearing a protective mask as a measure of protection against the *coronavirus COVID-19* (20-04-11 US) .

- Example 13: The Dutch government has ordered all schools and offices to be closed until further notice in attempt to control the spread of the *COVID-19 Coronavirus* (20-04-11 US).

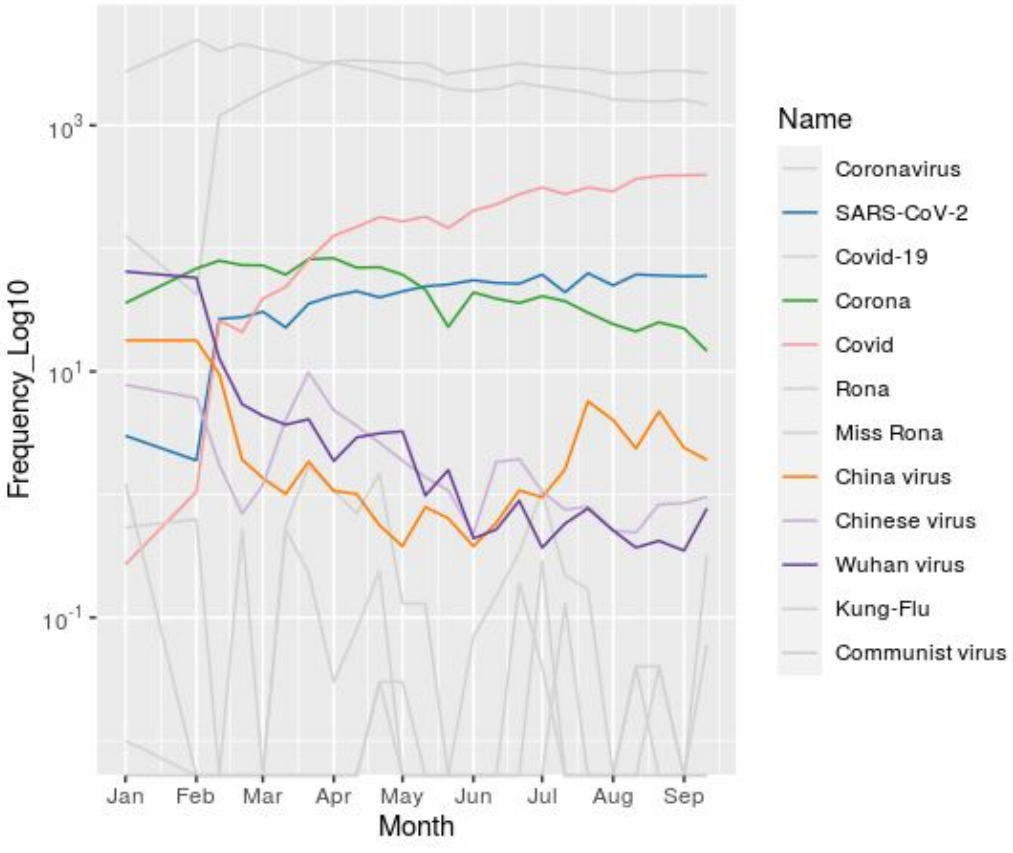
Whereas Covid becomes a widespread terms and needs no further clarification:

- Example 14: They're still seeing their rates of COVID patients go up exponentially (20-04-03 US).

In order to observe the least frequent terms, those that appear clumped together at the bottom of Figure 5, a logarithmic scale was applied to the Y axis. After, the names that underwent relevant changes were highlighted in order to better tell them apart from the rest. The result is plotted in Figure 6.

Figure 6

Most relevant changes in name popularity, highlighted. Log 10 scale on y axis.



Some interesting changes in name popularity are observable in Figure 6. First, we see a clearer picture of most popular names at the beginning as compared to the end. Wuhan-virus was, in January, almost as popular as Covid-19, and more common than Corona, but its use went down considerably from February onwards. China virus also decreased in February. SARS-CoV-2 started out very low, and then rose in February, going on to take over Corona in May.

February emerges as another key point in time for changes in Covid-19 nomenclature. When looking at concordance lines from the month, we find that the problematic implications of using geographically based names such as Wuhan virus or China virus were beginning to be discussed:

- Example 15: Today, the 2019-nCOV has already spawned a number of names that are inappropriate - like the "*Wuhan Virus"* using the city's name or the *"China virus"* (20-02-05 CA).

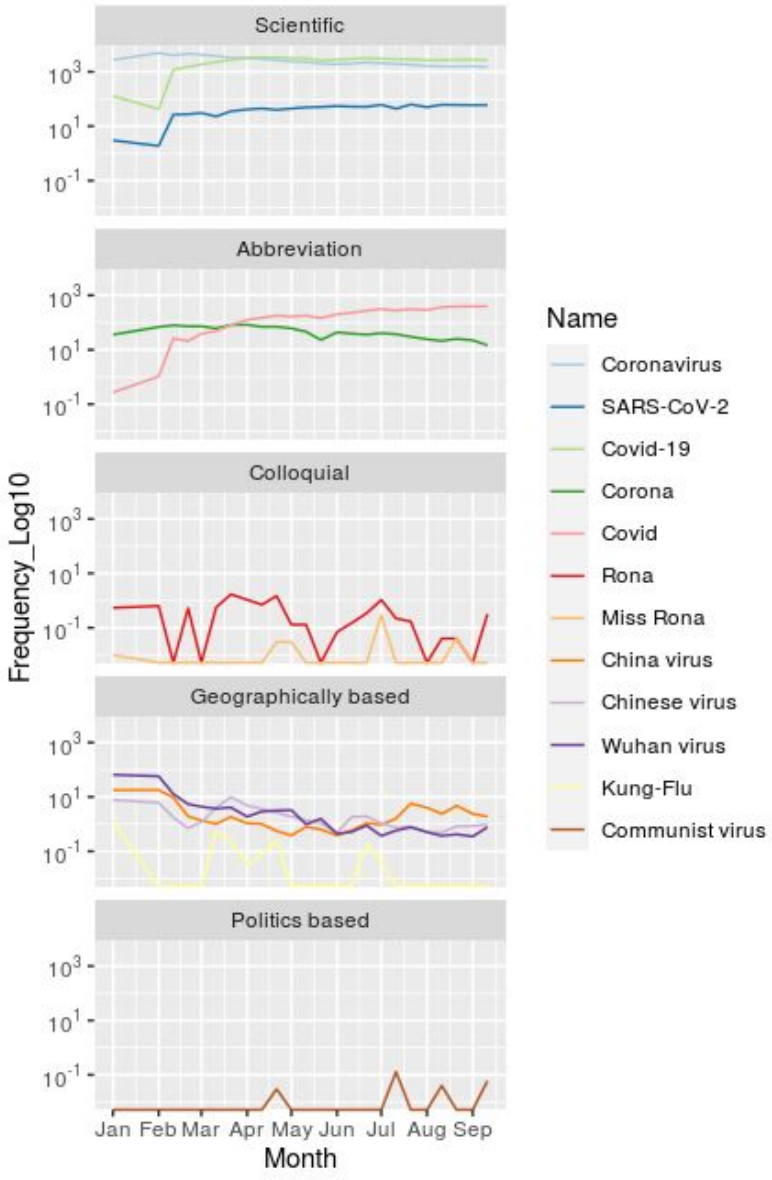
A for SARS-CoV-2, many articles make reference to the new 'proper' name for the disease:

- Example 16: An expert commission named the virus Sars-CoV-2 (20-02-15 NG).

In order to examine the remaining names, the chart in Figure 7 was designed, using the same logarithmic scale for the Y axis, but dividing them in five types (see Table 1).

Figure 7

Names for Covid-19 divided in types. Log 10 scale in Y axis.



For the scientific names, the slight uptick in the use of Covid-19 and SARS-CoV-2 after their use became widespread in February is visible. For the Abbreviation names, we find that Covid took over Corona in April, and kept on rising until the present day. The Colloquial names remain uncommon, although peaks in their use are visible, probably due to the high number of 0 frequencies in the data. For the Geographically based names, we can see a general decline in time, with only China Virus rising again after July, and Kung-Flu remaining very uncommon. Lastly, for the Politics based names, Communist virus is barely

used with the little peaks signifying every time it was used by a public figure, and the backlash that followed.

When looking at concordance lines for some of the names, a few interesting trends can be observed. First, Corona has a large array of alternative meanings and is not always used to refer to Covid-19:

- Example 17: Alexandria Ocasio-Cortez (D-Bronx), held a news conference in the *Corona* neighborhood of Queens (20-04-14 US).
- Example 18: But no worries about one of my fave concert venues, the *Corona* Theatre (20-04-15 US).

In addition, neologisms formed with Corona and Covid such as Corona Warriors or Covid Tuesdays were formed:

- Example 19: The Deck Down Under … serves a variety of sandwiches, pizza and seafood, while offering a "Surviving *COVID* Tuesdays" special (20-06-16 US).

And the personification of Corona emerges, with examples like:

- Example 20: *Corona* is the devil and it can not survive in the body of Jesus (20-05-03 ZA).

Last, Rona and Miss Rona remain uncommon and are usually attributed to younger speakers:

- Example 21: Miss Rona is taking too much light and some of us are left in the dark (20-04-23 CA).

5. Discussion

This chapter contextualizes and explains the results of the analysis conducted in chapter four in the context of the research presented in the Literature Review, as well as the socio-political circumstances of the Covid-19 timeline.

The first part of the analysis involved a comparison of Covid-19 and Wuhan virus regarding their use as names. Langendonck & Velde (2016) determined that the main characteristics of proper names are 'definiteness', 'unique denotation' and 'an inherent basic level sense" (35) to which Motschenbacher (2020) added a series of orthographic, morphological and syntactic characteristics, such as the prototypical absence of a determiner before a proper name. Through qualitative analysis, it was found that while the name Covid-19 fulfilled all of the characteristics outlined by the authors (particularly concerning capitalization and absence of a preceding determiner), Wuhan virus, in contrast, showed far more variation, with the words appearing both capitalized and in lowercase, often in quotation marks and preceded by the determiner 'the'. Since Covid-19 was introduced by WHO as a capitalized noun from the very beginning and Wuhan virus is a popular name propagated by the press, the differences in capitalization between the two cannot really be studied as organic and depending on the degree of familiarity, as originally stated by Langendonck & Velde (2016: 53). Rather, the reason for their differences could be that Covid-19 has more 'eponymous' characteristics, seeming more name-like than Wuhan virus, in which Wuhan could be interpreted as a modifier as in 'one virus originated in Wuhan, a Wuhan virus'. As for their definiteness, unique denotation and basic level sense, both names clearly refer to an unique entity and include the basic sense 'disease'.

The second part of the analysis concerned the normalized frequency counts collected for each of the twelve names studied. The greatest shifts in the popularity of names for Covid-19 can be found in February and April, while some of the names (i.e. Communist

virus) exist only in small peaks around extremely well-defined dates. In February, Wuhan virus and China virus went down in popularity, while SARS-CoV-2 rose. In April, Covid-19 took over Coronavirus and Covid also became more common. From March to May, peaks in the use of Chinese virus and Kung-Flu are visible and again from mid-June to mid-July. From April to May, July to August, and August to September small peaks are visible in the use of Communist virus.

When looking at the Covid-19 timeline (see section 2.1), one can see that on February the 11th the "WHO announced that the disease caused by the novel coronavirus would be named COVID-19", similarly, on March the 2nd the ICTV introduced the name SARS-CoV-2 (Gorbalenya et al. 2020). At the same time, concerns over the racist implications of geographically based names for Covid-19 were starting to arise (Forster 2020; Nassar & Kelly 2020; Lee 2020). In April, the name Covid-19 (sometimes reduced to Covid) had become widespread and was being used in all of WHO's reports to refer to the virus (Organization 2020c), while Coronavirus started being disregarded as a too general term. In the beginning of March, Trump and other White House officials used the term Chinese virus, attracting backlash (Itkowitz 2020), also in March, the term Kung-Flu was starting to trend on twitter (Butler 2020), which also caused considerable controversy. Therefore, there is a clear correlation between the timeline and the changes in popularity found through the analysis. Not only the actions of scientific organizations like WHO and public figures like President Trump directly impacted the use of one name or another for the virus, but – perhaps counter-intuitively , and against the perception of scholars like Rosen et.al. (2018) – names suggested by the scientific community according to the WHO and ICTV guidelines, when introduced, were preferred to 'catchy' abbreviations and popular names; and public backlash against racist geographically based names resulted in a considerable drop in their use.

However, the results gathered from this analysis cannot speak for the preferences of the general population, as Mark Davies' Coronavirus Corpus is a collection of articles published by online media sources, who are generally subjected to a higher level of scrutiny regarding integrity and ethics than the average person, particularly in an age in which 'cancel culture' is becoming more and more prominent. A corpus containing everyday language in addition to online news articles would have to be examined in order to confirm or disprove the results of the present study.

6. Conclusion

To conclude, this chapter briefly summarizes the results in the context of the research questions, as well as the initial working hypotheses to be addressed in the study. This study intended to serve as an exploration of the different names used for Covid-19 through time and answer the two research questions: (1) Whether terms used to name Covid-19 can be considered prototypical names and differentiated from common nouns according to the characteristics outlined by Onomastics; and (2) whether the popularity of the most salient names for Covid-19 has evolved through time, what type of names this evolution has favored, and how it relates to the socio-political circumstances of the Covid-19 timeline.

Through a comparison of Covid-19 and Wuhan virus, the study found that:

- The name Covid-19 largely fulfills the characteristics to be considered a proper name, being that it is always capitalized and never preceded by a determiner when used as a stand-alone term in the examples. Although harder to prove syntactically, the name appears to be definite, to have the inherent basic sense ‘disease’ and a unique denotation.
- The name Wuhan virus is a less prototypical name, it is not capitalized in some examples and it appears both in-between quotation marks and preceded by the antecedent ‘the’, although not ‘a’, showing a lack of contrastive definiteness. The name retains the inherent basic sense ‘disease,’ but it is less clear if it can be said that it has definitess and a unique denotation.

Examining the normalized frequency (PMW) for twelve of the names for Covid-19 across time, it was found that:

- The socio-political circumstances of the Covid-19 timeline (e.g. backlash against racism) as well as the language used by scientific organizations and public figures, like President Trump, had direct effects on the popularity of certain names for

Covid-19. In addition, the scientific names Coronavirus and Covid-19 were the most popular, surpassing abbreviations, geographically and politics-based names, and colloquial names. In fact, SARS-CoV-2, the most complex name, has risen in popularity in the recent months, going on to surpass the name Corona, contradicting the idea that names following the WHO guidelines will not be used due to being too complicated, counter-intuitive and not catchy enough.

Further research is needed in order to come to definite conclusions on the naming processes for Covid-19, as the corpus used in the present study − the Coronavirus Corpus by Mark Davies (2020) − only includes online news articles, and therefore cannot reflect the everyday language used by the general population.

References

Ainiala, Terhi. 2016. Names in Society. In Carole Hough (ed.), *The Oxford Handbook of Names and Naming*. Oxford: Oxford University Press.

Ainiala, Terhi & Jan-Ola Östman. 2017. *Socio-onomastics: The pragmatics of names*. Amsterdam: John Benjamins Publishing Company.

Burge, Tyler. 1973. Reference and Proper Names. *The Journal of Philosophy*. Journal of Philosophy, Inc. 70(14). 425–439. https://doi.org/10.2307/2025107 (6 September, 2020)..

Butler, Rhett. 2020. Twitter Explodes Over Calling Coronavirus "Kung Flu." *The Source*. https://thesource.com/2020/03/09/twitter-explodes-over-calling-coronavirus-kung-flu/ (2 September, 2020).

Callaway, Ewen & David Cyranoski. 2020. China coronavirus: Six questions scientists are asking. *Nature*. Nature Publishing Group 577(7792). 605–607. https://doi.org/10.1038/d41586-020-00166-6 (7 September, 2020).

Davies, Mark. 2016. *Corpus of News on the Web (NOW): 10 billion words from 20 countries, updated every day.* Available online at https://www.english-corpora.org/now/.

Davies, Mark. 2020. *The Coronavirus Corpus*. Available online at https://www.english-corpora.org/corona/.

Donnellan, Keith S. 1972. Proper Names and Identifying Descriptions. In Donald Davidson & Gilbert Harman (eds.), *Semantics of Natural Language*, 356–379. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-010-2557-7_10 (6 September, 2020).

Evans, Michael J. & Rainer Wimmer. 1990. Searle's theory of proper names, from a linguistic point of view. In Armin Burkhardt (ed.), *Speech Acts, Meaning and Intentions. Critical Approaches to the Philosophy of John R. Searle.*, 259–2784. Berlin: De Gruyter.

Forster, Victoria. 2020. Coronavirus Gets A New Name: COVID-19. Here's Why That Is Important. *Forbes*. https://www.forbes.com/sites/victoriaforster/2020/02/11/coronavirus-gets-a-new-name-covid-19-heres-why-renaming-it-is-important/ (2 September, 2020).

Gorbalenya, Alexander E., Susan C. Baker, Ralph S. Baric, Raoul J. de Groot, Christian Drosten, Anastasia A. Gulyaeva, Bart L. Haagmans, et al. 2020. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology* 5(4). 536–544. https://doi.org/10.1038/s41564-020-0695-z (5 September, 2020).

Henninger, Daniel. 2020. Opinion | A Communist Coronavirus. *Wall Street Journal*, sec. Opinion. https://www.wsj.com/articles/a-communist-coronavirus-11580341911 (3 September, 2020).

Hough, Carole. 2016. Introduction. In *The Oxford Handbook of Names and Naming*, 1–35. Oxford: Oxford University Press.

Hui, Mary. 2020. Why won't the WHO call the coronavirus by its name, SARS-CoV-2? *Quartz*. https://qz.com/1820422/coronavirus-why-wont-who-use-the-name-sars-cov-2/ (2 September, 2020).

Itkowitz, Colby. 2020. CDC director rejects label 'Chinese virus' after Trump, McCarthy tweets. *Washington Post*. https://www.washingtonpost.com/politics/cdc-director-rejects-label-chinese-virus-after-trump-mccarthy-tweets/2020/03/10/58bd086c-62e5-11ea-b3fc-7841686c5c57_story.html (2 September, 2020).

Jiang, Weija [@weija]. 2020. T*his morning a White House official referred to #Coronavirus as the "Kung-Flu" to my face. Makes me wonder what they're calling it behind my back.* [Tweet]. https://twitter.com/weijia/status/1239923246801334283 (3 September, 2020).

Jordan, Jerilyn. 2020. Ted Nugent Wants List of All U.S. Deaths From Past 5 Years to Show "Chinese Communist Virus" Numbers Are Fake. *San Antonio Current*, sec. Detroit Metro Times. https://www.sacurrent.com/sa-sound/archives/2020/08/28/ted-nugent-wants-list-of-all-us-deaths-from-past-5-years-to-show-chinese-communist-virus-numbers-are-fake (2 September, 2020).

Langendonck, Willy van & Mark van de Velde. 2016. Names and Grammar. In Carole Hough (ed.), *The Oxford Handbook of Names and Naming*, 35–56. Oxford: Oxford University Press.

Lee, Marie Myung-Ok. 2020. The racist art of naming a virus. *Salon*. https://www.salon.com/2020/02/06/the-racist-art-of-naming-a-virus/ (2 September, 2020).

Lesney, Mark. 2020. SARS-CoV-2: What's in a Name? *Medscape*. http://www.medscape.com/viewarticle/925710 (4 September, 2020).

Libin, Kevin. 2020. Kevin Libin: Beijing's communist regime is the biggest "China virus" threatening our survival. *National Post*. https://nationalpost.com/opinion/kevin-libin-beijings-communist-regime-is-the-biggest-china-virus-threatening-our-survival (3 September, 2020).

Long, Ralph B. 1969. The Grammar of English Proper Names. *Names*. Routledge 17(2). 107–126. https://doi.org/10.1179/nam.1969.17.2.107 (7 September, 2020).

Majewska, Ewa. 2014. Eponyme in der deutschsprachigen medizinischen Fachpresse. In C. Hough, & Izdebska, D. (eds.), *Names and Their Environment. Proceedings of the 25th International Congress of Onomastic Sciences, Glasgow, 25-29 August 2014.* 94-106. Glasgow: University of Glasgow. http://eprints.gla.ac.uk/122323/ (8 September, 2020).

Motschenbacher, Heiko. 2020. Corpus Linguistic Onomastics: A Plea for a Corpus-Based Investigation of Names. *Names* 68(2). 88–103. Abingdon: Routledge. https://doi.org/10.1080/00277738.2020.1731240 (8 September, 2020).

Nassar, Hana Mae & Ash Kelly. 2020. "Racist", "irresponsible": B.C. newspaper under fire for novel coronavirus-related headline. *News 1130*. https://www.citynews1130.com/2020/02/06/bc-newspaper-novel-coronavirus-headline (2 September, 2020).

Organization, World Health. 2020a. Novel Coronavirus (2019-nCoV): situation report - 1. *World Health Organization*. World Health Organization. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf?sfvrsn=20a99c10_4 (10 September, 2020).

Organization, World Health. 2020b. Novel Coronavirus (2019-nCoV): situation report - 10. *World Health Organization*. World Health Organization. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200130-sitrep-10-ncov.pdf?sfvrsn=d0b2e480_2 (10 September, 2020).

Organization, World Health. 2020c. Timeline: WHO's COVID-19 response. *World Health Organization*. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline (12 September, 2020).

Organization, World Health. 2020d. WHO Statement Regarding Cluster of Pneumonia Cases in Wuhan, China. *World Health Organization*. https://www.who.int/china/news/detail/09-01-2020-who-statement-regarding-cluster-of-pneumonia-cases-in-wuhan-china (12 September, 2020).

Rosen, Michael, Laura Wright, Laura Spinney & Peter Riot. Word of Mouth: Naming Diseases. https://www.bbc.co.uk/sounds/play/b09qd70t (23 August, 2020).

RStudio Team. 2016. *RStudio: Integrated Development for R. RStudio*. Boston: RStudio Inc. http://www.rstudio.com/.

Searle, John R. 1958. Proper Names. *Mind,* 67(266). 166–173. Oxford: Oxford University Press. https://www.jstor.org/stable/i313319 (12 September, 2020).

Shariatmadari, David. 2020. Pando, Miss Rona and Covid Toe: how the language of a disease develops – shaped by fear and prejudice. *The Guardian*. http://www.theguardian.com/world/2020/aug/20/pando-miss-rona-covid-toe-language-disease-fear-prejudice (3 September, 2020).

Shu, Lele. 2020. Avoid stigmatizing names for 2019 novel coronavirus. *Nature*. Nature Publishing Group 578(7795). 363–363. https://doi.org/10.1038/d41586-020-00458-x.

Sykes, William. 1896. On The Origin And History Of Some Disease-names. *The Lancet*, 147(3789), 1007–1010. https://doi.org/10.1016/S0140-6736(01)39505-3 (8 September, 2020).

Tse, Grace Y. W. 2003. Validating the Logistic Model of Article Usage Preceding Multi‑word Organization Names with the Aid of Computer Corpora. *Literary and Linguistic Computing*. Oxford Academic 18(3). 287–313. https://doi.org/10.1093/llc/18.3.287 (9 September, 2020).

Tse, Grace Y. W. 2004. A Grammatical Study of Personal Names in Present-Day English: with Special Reference to the Usage of the Definite Article. *English Studies*. Routledge 85(3). 241–259. https://doi.org/10.1080/00138380412331339140.

Tse, G. W. (2005). *A Corpus-based Study of Proper Names in Present-day English.* Bern, Switzerland: Peter Lang D. https://www.peterlang.com/view/title/46881 (13 September, 2020).

Tse, Grace Yuen Wah. 2002. *A corpus-based study of proper names in present-day English*. University of Lancaster Ph.D. https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.403781 (13 September, 2020).

Bredart, S., Brennen, T., Valentine, T. (1996). *The Cognitive Psychology of Proper Names.* London: Routledge, https://doi.org/10.4324/9780203132364 (14 September, 2020).

Wei, Fang, Yang Cheng & Cui Meng. 2020. Ringing the alarm. *Global Times*. https://www.globaltimes.cn/content/1178756.shtml (2 September, 2020).

Wong, Gary, Yu-Hai Bi, Qi-Hui Wang, Xin-Wen Chen, Zhi-Gang Zhang & Yong-Gang Yao. 2020. Zoonotic origins of human coronavirus 2019 (HCoV-19 / SARS-CoV-2): why is this work important? *Zoological Research* 41(3). 213–219. https://doi.org/10.24272/j.issn.2095-8137.2020.031 (14 September, 2020).