"If I'm contagious, I may infect other people": an anatomy of CONTAGION

Marco Bagli

University of Genoa
(marco.bagli@unige.it)

Abstract

Since its first identification in early 2020, the new coronavirus Sars-CoV-2 has quickly spread around the world, and the OMS declared the outbreak of a pandemic in March 2020. Since the early months of the pandemic, the attention of scientists, politicians and citizens has been drawn to the spreading of the virus, thus making the concept of contagion salient in daily news reports. In Italian, expressions such as '*il numero dei contagi*' became popular, as well as preoccupations regarding the number of *contagiati*, i.e., people who contracted the virus. Crucially, a direct translation of this lexical item into English is not possible, as it lacks the verbal lexical item in the semantic domain of *contagion.* Several dictionaries consistently report the verb "to infect" as a translation for It. *contagiare*, but specialised, medical terminology specifies a difference between the two conditions (i.e., an infection is different from a contagion).

1. Introduction

Since its first identification in the region of Wuhan, China, the virus Sars-CoV-2 has quickly spread around the world, and in March 2020 a pandemic was declared. The language of the pandemic has sparked a lot of interest among linguists. Many scholars have devoted their attention to the figurative conceptualisation of e.g., the virus as an *enemy* within the metaphorical frame of WAR[1] (e.g., Bagli 2021, Craig 2020, Sabucedo *et al*. 2020, Wicke and Bolognesi 2020, 2021). In reaction to the widespread usage of this metaphor, the #ReframeCovid project "was born as an open, collaborative and non-prescriptive initiative" (Olza *et al*., 2021) to gather linguistic and pictorial material that showcases alternative metaphorical realisations in any language. Some of the most fruitful and efficient alternative metaphors are the PANDEMIC IS A FIRE (Semino 2021), and at a more general level, the PANDEMIC IS A NATURAL FORCE. Other linguists have focused on Sentiment analysis of tweets, including an investigation of our understanding and response to Covid-19 (e.g., Chen *et al*. 2020, Combei and Luporini 2021, Wolohan 2020). Finally, several contributions have undertaken a more lexicographic approach, by mapping language change through the exploration of new lexical items (e.g., *covidiot, quarantini, blursday*) and the rapid

[1] In keeping with the typographical conventions of cognitive linguistics scholarship, I use SMALL CAPS to refer to concepts, while *italics* is used for lexical items.

acquisition of new meaning by conventional lexical material (e.g., *elbow-bump, front-lines*) (Kranert et al. 2020, Zheng 2020).

The present investigation is an exploratory study nestled in the lexicographic line of inquiry. It explores the realisations of the concept of CONTAGION in English. It aims at answering the following research questions: How does English lexically encode the concept of CONTAGION? What is the grammatical and semantic construal of the lexical items in this domain? What are the differences between the linguistic encoding of the two concepts between English and Italian? To answer these questions, I gathered linguistic data in a usage-based perspective, in keeping with cognitive linguistics commitments. The results of the analysis show a differential preference in the concepts of contagion and infection between the two languages.

The paper proceeds as follows. After this brief introduction, I review the major tenets and main ideas of cognitive linguistics that will shape the reasoning and the interpretation of results in the following sections. Paragraph 3 reviews the Research Questions and Methodology, while Paragraph 4 provides an overview of the data. Paragraph 5 defines the concept of Infection, while Paragraph 6 showcases the preference for the concept of contagion in an Italian corpus of Covid-related news articles (Busso and Tordini 2021). Finally, Paragraph 6 concludes the paper.

2. Theoretical background: Cognitive linguistics

Cognitive linguistics is a multifaceted approach to the scientific study of language that unites several strands of research with different foci. Scientists working in this tradition assume that our ability to produce language relies on more general cognitive skills, that language evolved to convey meaning, and that linguistic knowledge emerges from language use (Croft and Cruse 2004; Dąbrowska and Divjak 2015: 1). According to the first hypothesis, the human ability to produce language relies on fundamental mechanisms of non-linguistic reasoning and cognitive processing that had evolved in our lineage. For instance: memory, categorisation, and judgement (Tomasello 2008), to name but a few. On one hand, the study of human languages may allow us to grasp insights about the inner workings of the human mind; on the other, theories and phenomena observed in non-linguistic tasks are reflected in language. Thus, the lexical elements that emerge in language are understood as representing concepts in our minds, despite ongoing debate on the nature of this relationship (Speed, Vinson, and Vigliocco 2015).

According to the second hypothesis, any linguistic element conveys meaning, from phonological units to syntactic relations. Cognitive linguists agree that meaning is created through construal operations, or conceptualisation processes, and that it is structured by image-schemas, i.e., schematic representations of embodied experiences (Lakoff 1987). The major categories of construal operations identified in literature are Attention and Salience, Comparison, and Perspectivization; the most common image-schemas arise from our perception of space (e.g., up-down, front-back), containment (e.g., in-out, full-empty), force (e.g., balance, compulsion), among others (for a comprehensive list see Croft and Cruse 2004: 45).

The importance of construal is best exemplified when alternative expressions for the same situation are licensed by the same language. For instance, consider (1), adapted from Croft and Cruse (2004: 41):

(1) a. The leave**s** on the tree are beautiful.
b. The foliage-**∅** on the tree is beautiful.

Sentences in (1a-b) describe the same situation, and yet the choice of the grammatical construal is different. In (1a) the noun *leaf* is inflected in its plural form, thus evoking an image of a multitude of individual elements juxtaposed (the leaves). This is an example of the counting construction, which construes the entity as discrete (i.e., bounded) and heterogenous (Croft 2000). The noun *foliage* in (1b) instead is a mass-noun, and in English it cannot be inflected for number: the construal that emerges from this choice is of an unbounded and internally homogenous whole (Croft 2000). The two alternative construals are perhaps more evident when the same nominal stem may support both, as in (2), from Croft and Cruse (2004: 41):

(2) a. We have chocolate-**∅** for dessert.
b. We have chocolate**s** for dessert.

The utterance in (2a) refers to the substance 'chocolate', which in English is a mass-noun. As such, it is construed in its singular, uncountable grammatical form. Sentence (2b) sponsors the same noun inflected for number, therefore referring to discrete, individual objects covered with the same substance, which may or may not be filled with it.
Construal operations represent the basic mechanisms in the creation of meaning, which in turn is structured around even more basic entities in the human mind: i.e., image-schemas, which are defined by Johnson as:

> a recurrent pattern, shape, and regularity in, or of, these ongoing ordering activities. These patterns emerge as meaningful structures for us chiefly at the level of our bodily movements through space, or manipulation of objects, and our perceptual interactions (Johnson 1987: 29).

The perceptual information mediated by our bodies is abstracted and schematised, and eventually elaborated to structure non-bodily experience and abstract concepts via figurative processes, such as metaphor and metonymy (Johnson 1987; Talmy 1983). Despite arising from experience and bodily experience, image schemas differ from conventional images in two relevant aspects: they lack details, and they do not convey specific knowledge (Lakoff 1987: 453). For instance, consider the sentences in (3):

(3) a. Cherries are *red*.
b. *Red* covered the ice. (from Sandford 2021: 220)

The lexical item *red* in (3a) is construed as an attribute, i.e., a property that characterises a salient feature of an object (cherries in this case). Conversely, in (3b) the lexical item *red* is conceptualised as a substance with concrete physical dimensions and properties, to the extent that it can *cover* another substance (ice in this case), but it is not conceived as being an intrinsic property of that

substance. The two alternative construals of the same lexical item are based on two different image schemas: our understanding in (3a) emerges from the *part-whole* image schema, which structures our understanding of colour as a property (i.e., a part) of an object (i.e., a whole). The semantic interpretation of (3b) instead arises from the object image schema, which imposes a conceptualisation of *red* as a concrete, bounded object (Sandford 2021).

Crucially, there are other mechanisms that intervene in our creation of meaning: metaphor and metonymy. Despite their consideration in previous linguistic literature as literary tropes and "figures of speech", Lakoff and Johnson's milestone "Metaphors We Live By" (1980) has influenced more than a generation of scholars in linguistics, who have proved that conceptual metaphor and metonymy are fundamental mechanisms in our understanding of reality. Conceptual metaphors not only structure our understanding of reality, but they may even influence our behavioural responses to the external world (e.g., Gibbs 2005; Schubert 2005; Casasanto 2017; Winter and Matlock 2017). While most of the scientific production has concentrated on metaphors, research on conceptual metonymy is speedily catching up (e.g., Brdar 2017). The difference between the two mechanisms is perhaps best exemplified with a comparison proposed by Pérez-Sobrino (2017). Metaphors are like bridges that link two otherwise unrelated concepts, whereas metonymies are like icebergs, in which the emerged tip signals there is more just under the surface. That is: conceptual metaphors instantiate a relation of analogy between two concepts, which may be associated on the base of perceived similarity. The relationship between concepts in a metonymy instead is of contiguity: the two concepts are tied together because they belong to the same domain. Thus, while we formulate conceptual metaphors as for instance PANDEMIC IS WARFARE (where the pandemic is target domain B and warfare is source domain A); we formulate conceptual metonymies as providing mental access to a target concept. For instance, the expression *Paris was the first to declare war to the virus*, the item *Paris* is the lexical vehicle that provides access to the target concept *French government*. The most common metonymies may be formulated as CAUSE FOR RESULT or RESULT FOR CAUSE.

Finally, the last assumption in cognitive linguistics suggests that linguistic and grammatical knowledge emerges from its use, both at a diachronic and synchronic level. This implies that any analysis and theories of linguistic phenomena should be based on data observed in their natural occurrence, thus urging for a usage-based approach to the study of language. While the importance of actual data is widely recognized in contemporary research, back in the days of its first discussion this was a rather bold statement, especially if compared to the "armchair" method of previous theoretical frameworks (Dąbrowska and Divjak 2015: 1). Linguists working in cognitive linguistics collect data either from elicited tests or from large collections of texts, such as corpora, that allow for quantitative analysis and generalisations on linguistic patterns.

The analysis proposed in this paper is deeply rooted in a cognitive linguistics prespective to language description. It adopts the methodology of corpus analysis to ascertain the linguistic construal of the concept of CONTAGION, by investigating the syntactic and grammatical configuration of the lexical items in this domain that are available to English speakers. Furthermore, it contrasts the results of the analysis on contagion with numerical information on the frequency of lexical items associated to the domain of INFECTION, both in English and Italian.

The results show a differential preference for the two concepts across the two languages under scrutiny.

3. Research Questions and Methodology

The concept of CONTAGION has become extremely entrenched in the discourse about the Coronavirus.

To verify its linguistic conceptualisation in English, I relied on the Coronavirus Corpus (Davies 2020), an online Corpus specifically dedicated to the collection of linguistic material about Coronavirus. To compare it with the Italian conceptualisation of *contagio*, I relied on the corpus compiled by Busso and Tordini (2021) for Italian, which, despite being smaller, it offers a valuable resource for the study of the language of the pandemic in Italian. The methodology I adopted is that of corpus linguistics: frequency lexical items, collocations and manual analysis of Key Words In Context (KWIC). The research questions that I aim to answer are:

- what is the conceptualisation of contagion in English?
- what are the differences in construal between English and Italian?
- are there different syntactic preferences between the two languages in the encoding of the concept?
- what is the English equivalent of '*il numero dei contagi*'?

4. CONTAGION

The lexical item *contagion* derives from Latin *contangere*, which is a compound of *con* 'together' and *tangere* 'to touch', from PIE root **tag-*. Cognate words are *contact, contaminate, intact, integrate, tact, tactics, tangent,* and *tangible*, among others. The meaning of "communication of a disease" (OED *contagion*, n., 1) had already developed in Latin. The image schema that underlies the concept of contagion is 'touching', which in turn arises from the mundane and deeply embodied experience of two bodies touching each other. The touching of the two bodies allows the transmission of the pathogen, which metonymically grants us access to the entire event of the transmission of a disease.

The definition of *contagion* (n.) in the Oxford English Dictionary (hence, OED) includes several distinct meanings. The first entry reports "The communication of disease from body to body by contact direct or mediate" (OED *contagion* 1.a), thus referring to the actual event of transmission of a pathogen between two bodies. This meaning is extended to refer also to the "contagious quality or influence" (OED *contagion* 1.b). The second entry in the OED defines *contagion* as "a contagious disease or sickness; a plague or pestilence" (OED *contagion* 2). The relationship between meaning 1.a and 2 is metonymic: I will review it in detail in paragraph 4.3. The third meaning reported by the OED is also derived through metonymy: "The substance or principle by which a contagious disease is transmitted" (OED *contagion* 3.a) (see paragraph 4.3). The image schema at the basis of the concept of *contagion* may also be further extended and elaborated through conceptual metaphor: "hurtful, defiling, or corrupting contact; infecting influence" (OED *contagion* 4.a; see example 12), which can also be specific to morality: "contagious or spreading

moral disease; moral corruption" (OED 4.b). Notably, *contagion* does not have necessarily a negative connotation. It may also refer to "the contagious or 'catching' influence or operation of example, sympathy, and the like" (OED 5), thus imposing an emotionally positive semantic construal on the lexical item.

4.1 Lexical items

The lexical items that pertain to contagion and that are present in the Coronavirus corpus are listed in Table 1. I retrieved them by searching the corpus with the wildcard *contag**. In Table 1, the first column refers to the lexical item, POS stands for Part of Speech, and the last column displays Frequency.

item	POS	FREQ.
contagious (also includes *contageous*)	adj.	18228
contagion	n.	12917
contagions	n.	478
contagiousness	n.	407
contagiously	adv.	22

Table 1. Lexical items in the domain of contagion.

Table 1 reports the lexical items retrieved by the search *contag** with their frequencies. The adjective *contagious* is the most frequent (Frequency= 18206). The search also yielded the item *contageous* (F= 22), which represents a spelling mistake of the most common *contagious* and therefore was added to the total Frequency of *contagious* in Table 1. The noun *contagion* (F= 12917) may also be inflected by number (*contagions,* F= 478), which however is dramatically less frequent than the other items. Similarly, the noun *contagiousness* (the status of being *contagious,* F= 407), derived from the adjective and the nominal suffix *-ness*, has a significantly lower frequency in the corpus than other lexical items. Finally, the adverb *contagiously* (F= 22) is rarely used.
Both the adjective *contagious* and the noun *contagion* originated in Latin and entered Middle English via French (OED, *contagious*, adj. and *contagion*, n.); the noun *contagiousness* instead originated in Early Modern English (OED, *contagiousness*, n., first attestation 1530).

4.2. Collocations

The Coronavirus corpus is part of the larger family of English Corpora, which includes COCA (Corpus of Contemporary American), COHA (Corpus of Historical American), and the iWeb corpus, among others. The user-friendly, online interface allows for the research of collocations within the corpus. Table 2 reports the ten most frequent collocations in the corpus for the lexical item *contagious.*

item	Frequency[2]
more	8280
highly	7396
virus	4819
disease	3317
variant	3060
variants	1900
spread	1814
strain	1155
diseases	974
deadly	764

Table 2. Collocations for contagious

The two most frequent collocates of the adjective *contagious* are the adverbs *more* and *highly.* The lexical item *deadly* is the only adjective in the list, and the other items are all nouns: these are *virus*, *disease*, *variant*, *variants*, *spread*, *strain,* and *diseases*. Overall, the list of collocations reflects the most common preoccupations related to the contagion during the ongoing pandemic.
Particularly, the two items *variant* and *variants* refer to the discourse around the mutations of Sars-Cov-2, which keeps evolving into new and different forms, often discussed with reference to their *contagiousness*:

(4) The Delta variant is much more contagious than previous variants we've seen. (31/08/21, Greenwich Time)

The collocations of the item *contagion* are reported in Table 2. The most frequent collocate is *risk*, followed by *spread*, *fear*, *fears*, *prevent*, *deadly*, *avoid*, *contain* and *risks.* The items unanimously reflect the danger represented by the contagion, the *risk* of its *spread*, the *fear* that it evokes in humans, and the need to halt it.

item	FREQ.
risk	871
spread	824
fear	398
fears	304
prevent	285
deadly	234
avoid	232
reduce	203
contain	201
risks	199

Table 3. Collocations of contagion

[2] The figures reported in this paper were extracted from the corpus in Spring 2021. At a later inspection, these have increased significantly, although the lexical items have remained the same. The main difference is the presence of the noun Delta, which identifies a variant of the virus that became prominent in public discourse during Summer 2021.

The contagion may be conceptualised as a substance that covers a specific area, as in (5-6):

(5) **spread** of the deadly/viral/lethal contagion
(6) the contagion **spread** as fast as the fear of death

The lexical item *spread* may be used either as a verb whose subject is *contagion*, or as a deverbal noun. The adjectives that most commonly fill the construction "spread of the ADJ contagion" are exemplified in (5), and they convey the danger of exposure to the virus.
The contagion may also be construed as an external phenomenon with its own propelling force, which needs to be *prevented*, *reduced*, and *contained* (7-9), thus conceptualising it as an opponent that needs to be defeated, and against which one needs to protect themselves.

(7) […] broad restrictions to **prevent** the spread of contagion.
(8) […] measures to **reduce** contagion.
(9) […] widespread lockdowns to **contain** contagion.

These construals are in keeping with the dominant metaphorical frames in the discourse of the current pandemic, namely the PANDEMIC IS WARFARE and the PANDEMIC IS A NATURAL FORCE.
Other collocates convey the feelings of humans towards the contagion, as in (10-11):

(10) […] **fear** of global/Covid-19 contagion
(11) […] **fears** over coronavirus contagion

Notably, the lexical item *fear* may be the object being transmitted through contagion, as opposed to being a reaction to the contagion:

(12) […] the contagion of **fear** infected markets

In (12), *fear* is metaphorically conceptualised as a pathogen that may cause an infection in the specific context of economic finance. The metaphorical understanding of the phrase in (12) is licensed by the figurative meaning of *contagion* (OED, *contagion*, 4.a). Thus, the concept of *contagion* may be successfully employed not only to describe the physical, concrete transmission of an organism between two bodies, but it may also be metaphorically extended to refer to any type of influence of one entity on another.
The verbs that most frequently collocate with *contagion* are reported in Table 4, and they either conceptualise the transmission of the disease as an ongoing process (*spreads*, *spread*, *spreading*, *continues*) or they focus on the initial moment of the event (*came*, *emerged*, *coming*, *became*, *resulting*, *started*).

item	FREQ.
came	51
spreads	41
continues	38
emerged	36
coming	32

spread	23
spreading	22
became	15
resulting	13
started	13

Table 4. Concordance of contagion + VERB

Overall, contagion is not frequently used as a syntactic subject. It is more frequently found as object of the preposition *of* in NPs headed by deverbal nouns such as *the spread of the contagion* (5, 7). The phrases in (13) report an example in which *contagion* is the subject of the verb *to emerge*:

(13) […] the largest daily rise since the contagion **emerged**.

The usage of the verb *to emerge* licenses a conceptualisation of the contagion as a phenomenon beyond human control and agency. The phrase in (13) displays a metonymic conceptualisation of the term *contagion*, which provides access to the entire event of the pandemic via the metonymy CAUSE FOR RESULT. Finally, the expression *daily rise* in (13) refers to the number of new cases of individuals contracting the virus.
Lastly, I checked the collocations for the lexical item *contagions* (Table 5).

item	FREQ.
other	48
spread	30
future	17
against	16
grow	14
prepared	14
wave	14
quickly	14
limit	12
increase	12

Table 5 Collocations of contagions

Despite being less frequent than its singular form (see Table 1), the analysis of the collocations of *contagions* offers relevant insights to the general construal of the concept. The lexical items that most frequently collocate with *contagions* are *other, spread, future, against, grow, prepared, wave, quickly, limit,* and *increase*. The verb *to spread* also collocates with the singular form of the noun (Table 3). The preposition *against* and the verb *to limit* license a construal of *contagions* as an opponent, which is reminiscent of the conceptualisation of the singular form, while the lexical item *wave* suggests a construal of the contagion as a large body of water. The same metaphorical frame has been observed for the pandemic, whose different phases have largely been referred to as *waves* (Semino 2021). A few occurrences suggest a conceptualisation of *contagion* in its second meaning (OED *contagion* 2):

(14) Colds and the flu are viral **contagions** that are spread through aerosols, just like Covid-19

The sentence reported in (14) displays an instance of the noun *contagions* used to refer to an entire epidemic, resulting from a metonymic construal in which the cause of the event (i.e., the contagion) provides mental access to its result (i.e., an epidemic/ illness). Unlike Italian *contagi*, the plural form of the noun in (14) does not refer to the collection of individual cases of disease transmission, rather it conceptualises the event of the contagion as a discrete, internally homogenous process.
There are however other collocates which license a construal of *contagions* as a series of individual transmission:

(15) The country has managed to slow down the spread of coronavirus but should be prepared for **contagions** to **grow** quickly.

The utterance in (15) showcases the conceptualisations of CONTAGION as referring to the transmission between individuals, and its plural form refers to the collection of single individual events. To verify the distribution of the two distinct meanings of the lexical item *contagion*, I analysed manually the occurrences of its plural form. An overview is offered in Table 6.

item	Frequency	disease (%)	individual (%)
other	48	35 (73%)	13 (27%)
spread	30	13 (43%)	17 (57%)
future	17	16 (94%)	1 (6%)
against	16	13 (81%)	3 (19%)
wave	14	0	14 (100%)
limit	12	0	12 (100%)
prepared/ grow/ quickly	9	0	9 (100%)
increase	8	0	8 (100%)
total	154	77 (50%)	77 (50%)

Table 6 The different facets of contagions

Table (6) reports the frequencies of the different facets of *contagions*. Incidentally, the two alternative meanings have the same frequency. The items that most commonly collocate with the "disease" meaning are *other*, *future,* and *against*, while the items that collocate with the "individual" meaning are *wave, limit, prepared/ grow/ quickly, increase*. The lexical item *spread* collocates almost equally with both meanings. The examples in (16) contrast the two distinct meaning, that collocate with the same lexical item:

(16) (a) If you are with **other** people, **contagions** outdoors are also unlikely, especially under sunlight.
(b) these interventions alone did not tame cholera, malaria, and **other contagions** that plagued Western societies.

The two different meanings of the item *contagion* are exemplified in (16a, b), and display different syntactic configuration. In (16b) the item *other* modifies *contagions*, thus imposing a bounded construal in keeping with meaning 2 reported by the OED. Sentence (16a) instead refers to a collection of individual episodes of viral transmission between individuals.

There is a third meaning of the item *contagion* that is registered in the OED, and that is exemplified in (17):

(17) Viruses, bacteria and other **contagions** are a fact of nature and no one is responsible for protecting you from nature.

The meaning reported in (17) corresponds to meaning 3a in the OED, and it is motivated by the metonymic construal RESULT FOR CAUSE, in which the CAUSE of a contagion (viruses and bacteria) are mentally accessed via the RESULT of their transmission (the contagion itself)[3].

4.3 Discussion

The concept of contagion is defined in the OED as having three distinct meanings (see paragraph 3). Its first meaning is modelled on the base of the image schema of 'touching', and it refers to the transmission of either concrete or abstract entities between two individuals. This specific transmitting event is commonly used in reference to pathogen and external agents, that are likely to cause an infection in the affected individual.

The second meaning of the lexical item refers to an entire disease, such as flu, cholera, etc. This meaning is construed as a metonymic elaboration on the concept of contagion, in which the physical act of transmission serves as the linguistic item that provides access to the adjacent domain of disease. The relationship between the two domains is CAUSE-RESULT, in which the *contagion* is the CAUSE of the disease.

Lastly, the third meaning of *contagion* is construed through the reverse metonymy of meaning 2. In this case, the lexical item *contagion* provides mental access to the bacteria, viruses, and other entities that are the CAUSE of the contagion. Thus, we may successfully analyse meaning 3 as the result of EFFECT FOR CAUSE metonymy.

Although the concept of contagion may be used to refer to individual transmission, the occurrences retrieved in the corpus for the noun inflected in the plural are equally split between meaning 1.a (individual transmission of the disease) and meaning 2 (disease). During the pandemic, the concept of contagion has been discussed repeatedly, especially with reference to the rates of individual cases and transmission of the virus. The low frequency figures for the item *contagions* suggest that this is not the preferred concept to convey information about the rates of individual transmissions.

5. Infection

The concept of INFECTION is intimately connected to the concept of CONTAGION, albeit there are some differences. Crucially, English-Italian dictionaries unanimously report the verb *to infect* as equivalent to It. *contagiare*. According to the OED, the verb *to infect* derives from Latin *inficere* 'to dye, to stain, to impregnate, to imbue, to taint, to poison, to affect with disease' (OED, *infect, v.*).

[3] This was the only case of meaning 3a that I retrieved from the corpus, and therefore I considered it within the "disease" category in Table 6.

The difference between the two verbs (and concepts) is subtle, but substantial. According to Merriam-Webster dictionary forum:

> *Contagious diseases are spread by contact, while infectious diseases are spread by infectious agents. Something "contagious" is by default "infectious" because contact exposed you to the infectious agent, but something infectious isn't always contagious. (https://www.merriam-webster.com/words-at-play/gesundheit-is-that-cold-contagious-or-infectious, last accessed 4 December 2021).*

There are infections that are not contagious, e.g., food poisoning caused by bacteria in the food ingested, which however do not spread among people by exposure to the infected individual. Notably, the concept conveyed by *to infect* displays two different facets: (a) to transmit a disease and (b) to cause a disease; conversely, *contagion* only describes the transmission of a pathogen through touch, and it implies the development of an infection. Despite there being possible *contagious* agents that do not develop an infection (e.g., *lice* and other parasites, https://www.rchsd.org/health-articles/head-lice/), an informal analysis on COCA suggests that the items lexically related to *contagion* do not frequently collocate with *lice*[4].

I searched the Coronavirus corpus with the query "*infect**" to retrieve the lexical items in the domain of infection. Table 7 reports the part of speech (POS) and Frequencies of the results of the query in the corpus:

item	POS	Frequency
infection	n.	202810
infections	n.	186322
infected	pp./adj.	185517
infectious	adj.	78381
infect(s)	v.	16490

*Table 7 Frequency of lexical item for infect**

As expected, the list encompasses a verbal element *to infect*, which however is not the most frequent item in the domain (F= 16490). The most frequent item instead is the noun *infection* and its plural form *infections*, which together represent 58% of the total occurrences of the lexical items. The comparison of the frequencies of this lexical item with those of *contagion* (Table 1) reveal that the domain of INFECTION is extremely more frequent than CONTAGION in the discourse of the pandemic. A close scrutiny of the occurrences reveal that the concept of INFECTION is indeed the preferred one to describe individual transmission of the pathogen, as exemplified in the utterances in (18-20):

(18) Now you are considered to have had a «close contact» with an **infected individual** if you've spent a cumulative 15 minutes over a 24-hour period.
(19) Every day we report a large number of people newly **infected** with COVID-19 […]
(20) The number of new Covid-19 **infections** has "consistenly increased over the past week"

[4] This claim stems from an informal analysis of the list of collocations of the lexical item *lice* in COCA, in which there is no sign of lexical items related to contagion. Lice however *infest*, *spread*, may be *picked* or *transmitted*.

The utterances in (18) and (19) display the past participle of the verb *to infect* to convey information on individual cases of disease-transmission, thus representing the English equivalent of the Italian *contagiati*. The utterance in (20) instead displays the plural noun *infections* to discuss the daily rise of new cases of Covid-19, and therefore is the equivalent of the Italian *il numero di (nuovi) contagi*. Although the two concepts are not strictly the same in scientific medical terminology, English and Italian systematically select one of the two terms to refer to the same event. In a cognitive linguistic perspective, the lexical choice of English may be successfully analysed as focusing on the *result* of the event, as opposed to the Italian preferred item, which instead focuses on the *cause* of the event. It must be stressed however that the two languages do not exclude alternative realisations.

6. A comparison with Italian

Italian operates different lexical choices in the conceptualisation and discussion of Coronavirus, and it systematically prefers the concept of contagion to talk about the transmission of the virus, while selecting *infection* less frequently. To verify this claim, I turned to the Italian covid corpus, compiled by Busso and Tordini (2021). This corpus was compiled by collecting newspaper magazines in a period between February 24th, 2020, and June 3rd, 2020. The two authors integrated Google searches (both manual and automatic) with the results of the web-scraping software BootCat (Baroni and Bernardini 2005). The Italian corpus is not fully comparable to the Coronavirus corpus, as it contains 362,464 tokens. It was designed avoiding daily reports but including interviews, investigative reports, and authoritative comments (Busso and Tordini 2021: 47-48). Nonetheless, the Italian Covid-corpus is representative of the Italian discourse about Covid-19, especially in the early weeks since the start of the pandemic.

I ran queries on the corpus through the software SketchEngine to retrieve linguistic data in a usage-based perspective. This allows me to avoid personal intuitions and introspection as a native speaker, in keeping with the cognitive linguistics perspective (Dąbrowska 2016). Considering the different nature of the two corpora, and the focus of my research being on English, I use the Italian corpus as a reference to retrieve (mainly) qualitative data and example. I report numerical figures for the frequencies of the lexical items retrieved in Italian: their relative frequencies within the same corpus are indicative of the lexical choices operated by the two languages.

Table 8 reports the frequencies of the lexical items pertaining to the concept of CONTAGIO in Italian. I aggregated the frequencies of inflected adjective, verb, and past participle. I kept the noun *contagio* separated from its plural form *contagi*, in keeping with the previous investigation on English. Furthermore, I considered the past participle form *contagiati* as a distinct form from the verb *contagiare*, on the base of its frequent usage as a nominalised, deverbal adjective, as (22) shows. The list of individual lexical items considered in the lemmas is provided in the footnote[5].

lemma	POS	

[5] The entry *contagiato* includes *contagiata, contagiati, contagiate; contagioso* includes *contagiosa, contagiose, contagiosi; contagiare* includes *contagia, contagiano, contagiando, contagianti*.

contagio	n.	240
contagi	n.	167
contagiato	p. participle	132
contagioso	adj.	21
contagiare	v.	15
contagiosità	n.	6

Table 8 Frequency of lexical items in contagio

From a comparison with the relative Frequency of English lexical items (Table 1), it clearly appears a difference in the POS preferred to talk about the contagion. While the noun *contagio* emerges as the most used in the Italian corpus, English prefers the adjective *contagious*, thus more frequently construing the concept of contagion as a property of an individual as opposed to a process and an event. The preference for these two different parts of speech by the two languages is related to the emergence of the verb *contagiare* in Italian. The grammatical categories of nouns and verbs are strictly connected from a syntactic point of view, to the extent that some linguists have described the Noun-Verb continuum (Simone 2020; Ross 1972).
In keeping with the expectations, the lemma *contagiato* is the most frequent after the noun (both in singular and plural). This lexical item is not represented in English, as it lacks the verbal element referring to the concept of CONTAGION, which despite being infrequent, is also present in the list of lexical items. Finally, the noun *contagiosità* refers to the status of being *contagious* (see *contagiousness*).

(21) […] portando così il numero complessivo dei **contagiati** a 67.366.

The utterance in (21) reports the nominalised, deverbal adjective of the verb *contagiare* in reference to individuals who contracted the virus, thus evoking a construal of the contagion as an individual event. The inflection for number in (21) describes a multitude of individual cases, thus evoking a construal of a bounded, internally heterogenous whole.
A manual analysis of the plural noun *contagi* suggests that this lexical item may also be used to conceptualise the individual event of transmission of the disease (23):

(22) A fronte di 1.116 tamponi, sono quattro i **contagi** scoperti dal sistema sanitario umbro.

The utterance in (22) displays the same usage of the noun inflected for number that was displayed in (16a). Interestingly, while there are no cases in the corpus of usage of the plural noun as in (16b) (i.e.: referring to "disease"), there are examples in which this meaning is realised with the singular noun, as in (23):

(23) Le regole di restrizione per limitare il contagio sono "giustificate" per il 75% degli italiani.

Finally, Table 9 reports the frequency of lexical items in INFEZIONE. In keeping with the organisation of the lexical items in Table 8, I grouped inflected lexical items per lemma, but I kept the nouns and the past participle separated. It emerges that the concept of INFEZIONE in Italian

has lower frequencies in its lexical representation in the corpus under scrutiny, thus suggesting that it is not the preferred concept to communicate the spreading of the virus.

item	POS	Frequency
infezione	n.	54
infettivo	adj.	45
infetto	adj.	40
infettato	p. participle	22
infezioni	n.	17
infettare	v.	14

Table 9 Lexical items of infezione

The lexical items related to the concept of INFEZIONE in Italian display richer morphology, including two different adjectives (*infetto* and *infettivo*) that evoke two distinct construals. The adjective *infetto* (i.e., infected) describes the status of something or someone who has been infected by something or someone, therefore conveying a passive meaning. Whereas the adjective *infettivo* (i.e., infectious) refers to the property of something to infect something or someone, thus conceptualising this adjective as an active property of someone or something. Notably, the status of "being infected" is also lexicalised by the past participle *infettato*, which evokes yet a different construal, in keeping with its syntactic category of verb. The item *infettato* highlights the passive meaning of the concept, suggesting that someone has been *infettato* by something else (as opposed to *infetto*, which instead denotes a more stable and permanent situation of being infected).

7. Conclusions

The present paper has explored the concept of CONTAGION, which has unfortunately become salient during the current pandemic caused by Sars-Cov-2. The lexical items that describe it originated in Latin and spread through the European languages, including English. Despite their common linguistic ancestor, the English and Italian concept display different construals, which in turn are reflected in different syntactic choices on the part of the two languages. Perhaps more strikingly, the two languages also differ in the preferred choice of concept to describe the same event, namely the collection of cases of individual transmissions of the virus. While Italian prefers the concept of CONTAGION, English prefers INFECTION. I argue that the reason for this preference lies in the different lexical items that are available to speakers of the language. While Italian includes the verb *contagiare* in its lexical repertoire, which licenses the passive construal of *contagiato* using the past participle, English lacks such possibility and relies on the verb *to infect* instead. The preference for the two different concepts predicts that the Italian expression *il numero dei contagi* should be translated in English as *the number of infected individuals/people*. Nonetheless, results of the corpus analysis have shown that alternative construals in the two languages are possible, but less common. The English sentence *the number of contagions* is grammatical (and retrieved in corpora), but it is not as frequent as *the number of infections*. These observations call for caution in the translation process and attest to the relevance of usage-based approaches to language description. Furthermore, the results suggest that Frequency of appearance in a corpus,

or lack thereof, is a paramount dimension that should not be underestimated in the description, production, and translation process.

Acknowledgements

The Italian Coronavirus corpus was kindly provided by Dr. Lucia Busso and Dr. Ottavia Tordini: thank you. I would also like to thank Dr. Fabio Ciambella for inviting me to the workshop "CONTAct, CONTAmination, CONTAgion in English linguistic, literary and cultural communication" organised by TrAdE research group at the University of Rome Tor Vergata in Spring 2021.

References

Bagli M., "The conceptualization of heroes during Covid-19 pandemic", in L. Busso and O. Tordini (eds), *Describing the new pandemic: linguistic and communicative studies on the COVID-19 crises*, Rassegna Italiana di Linguistica Applicata (RILA), LII (1-2), 2021, pp. 67-90.

Baroni M, and Bernardini S., "BootCaT: Bootstrapping corpora and terms from the web", in M.T. Lino *et al.*, *Proceedings of Fourth International Conference on Language Resources and Evaluation* (LREC 2004), Lisbon, ELDA, 2004, pp. 1313-1316.

Busso L., and Tordini O., "'Ricostruire senza Macerie'. La crisi sanitaria da Covid-19 nel linguaggio giornalistico italiano tra Fase 1 e Fase 2", in L. Busso and O. Tordini (eds), *Describing the new pandemic: linguistic and communicative studies on the COVID-19 crises*, Rassegna Italiana di Linguistica Applicata (RILA), LII (1-2), 2021, pp. 91-114.

Brdar M., *Metonymy and Word-Formation. Their Interactions and Complementation*, Newcastle upon Tyne, Cambridge Scholars Publishing, 2017.

Casasanto D., "The hierarchical structure of mental metaphors", in B. Hampe (edited by), *Metaphor: Embodied cognition and discourse*, Cambridge, Cambridge University Press, 2017, pp. 46-61.

Chen E. *et al.*, "Tracking social media discourse about the Covid-19 pandemic: development of a Public Coronavirus Twitter data set", *JMIR Public Health and Surveillance*, 6 (2), e19273, 2020. doi: 10.2196/19273.

Combei C.R. and Luporini A., "Sentiment and Emotion analysis meet appraisal. A corpus study of tweets related to the Covid-19 pandemic", in L. Busso and O. Tordini (eds), *Describing the new pandemic: linguistic and communicative studies on the COVID-19 crises*, Rassegna Italiana di Linguistica Applicata (RILA), LII (1-2), 2021, pp. 91-114.

Craig D., "Pandemic and its metaphors: Sontag revisited in the Covid-19 Era", *European Journal of Cultural Studies*, 23 (3), 2020, pp. 1-8. https://doi.org/10.1177/1367549420938403

Croft W., Countability in English Nouns denoting physical entities: a Radical Construction Grammar analysis, 2000, available online at https://www.unm.edu/~wcroft/Papers/Count.pdf

Croft W., and Alan Cruse D., *Cognitive linguistics*, Cambridge, Cambridge University Press, 2004.

Dąbrowska E., and Divjak D., "Introduction", in E. Dąbrowska and D. Divjak (eds), *Handbook of Cognitive Linguistics*, Berlin, De Gruyter Mouton, 2015, pp. 1-9.

Dąbrowska E., "Cognitive Linguistics' seven deadly sins", *Cognitive Linguistics*, 27 (4), 2016, pp. 479-491. https://doi.org/10.1515/cog-2016-0059

Davies M., *Coronavirus Corpus*, 2020, available at https://www.english-corpora.org/corona/

Gibbs R.W. Jr., *Embodiment and cognitive science,* Cambridge, Cambridge University Press, 2005.

Johnson M., *The body in the Mind: The Bodily Basis of Meaning, Imagination and Reason*, Chicago, University of Chicago Press, 1987.

Kranert M. *et al*., "Covid-19: The world and the words. Linguistic means and discursive constructions", in *Discourse Net Collaborative Working Paper Series,* Special Issue: Discourse Studies Essays on the Corona-Crisis, 2(9), 2020, pp. 1-7.

Lakoff G., *Women, Fire and Dangerous Things*, Chicago, University of Chicago Press, 1987.

Lakoff G., and Johnson M., *Metaphors We Live By*, Chicago, University of Chicago Press, 1980.

Olza I. *et al*., "The #ReframeCovid Initiative", in *Metaphor and the Social World*, 11(1), 2021, pp. 98-120. https://doi.org/10.1075/msw.00013.olz

Pérez-Sobrino P., *Multimodal Metaphor and Metonymy in Advertising*, Amsterdam-Philadelphia, John Benjamins, 2017.

Ross J.R., "The Category Squish: Endstation Hauptwort", *Chicago Linguistics Society*, 8, 1972, pp. 316-328.

Sabucedo J.M. *et al*., "Covid-19 and the metaphor of war (Covid-19 y la metáfora de la guerra)", *International Journal of Social Psychology,* 35(3), 2020, pp. 618-624.

Sandford J.L., *The Sense of Color. A Cognitive Linguistics analysis of color words*, Perugia, Aguaplano, 2021 (*in press*).

Semino E. "'Not Soldiers but Fire-fighters' – Metaphors and Covid-19", *Health Communication,* 36 (1), 2021, pp. 50-58. https://doi.org/10.1080/10410236.2020.1844989

Schubert T.W., "Your Highness: Vertical Position as Perceptual Symbols of Power", *Journal of Personality and Social Psychology*, 89 (1), 2005, pp. 1-21.

Simone R., *Il software del linguaggio*, Milano, Raffaello Cortina Editore, 2020.

Speed L.J., Vinson D.P. & Vigliocco G., "Representing Meaning", in E. Dąbrowska and D. Divjak (eds), *Handbook of Cognitive Linguistics*, Berlin, De Gruyter Mouton, 2015, pp. 190-211.

Talmy L., "How language structures space", in H.L. Pick Jr., and L.P. Acredolo (eds), *Spatial orientation: theory, research and application*, New York, Plenum Press, 1983, pp. 225-282.

Tomasello M., *Origins of Human Communication*, Cambridge, The MIT Press, 2008.

Wicke P., and Bolognesi M., "Framing Covid-19: How we conceptualize and discuss the pandemic on Twitter", in *PLoS ONE*, 15(9), 2020, p. e0240010. https://doi.org/10.371/journal.pone.0240010

Wicke P., and Bolognesi M., "Covid-19 Discourse on Twitter: How the Topics, Sentiments, Subjectivity, and Figurative Frames Changed Over Time", *Frontiers in Communication*, 6, 2021. DOI=10.3389/fcomm.2021.651997

Winter B., and Matlock T., "Primary metaphors are both cultural and embodied", in Beate Hampe (edited by), *Metaphor: Embodied cognition and discourse*, Cambridge, Cambridge University Press, 2017, pp. 99–115.

Wolohan J.T., "Estimating the effect of Covid-19 on mental health: linguistic indicators of depression during a global pandemic", in *Proceedings of the NLP-Covid-19 workshop*, Seattle, ACL, 2020.

Zheng Y., "Mobilizing foreign languages students for multilingual crisis translation in Shanghai", *Multilingua*, 39 (5), 2020. https://doi.org/10.1515/multi-2020-0095