

Judit Papp

How the COVID-19 pandemic is changing the Hungarian language: Building a domain-specific Hungarian/Italian/English dictionary of the COVID-19 pandemic

1 Introduction

This paper presents the main issues connected with the creation of a trilingual Hungarian-Italian-English dictionary of the COVID-19 pandemic using Lexonomy.¹ My aim is not only to create a coronacorpus (in Hungarian, I propose my own corona-neologism or ‘coroneologism’:² *koronakorpusz*) and a dictionary of equivalents, but also to understand how the different waves and phases of the COVID-19 pandemic are changing the Hungarian language, detect the *Corona-*, *COVID-*, *pandemic-*, *virus-*, *mask-*, *quarantine-*, and *vaccine-*related neologisms, and offer an overview of the most frequent or linguistically interesting Hungarian neologisms and multiword units related to COVID-19.

For the creation of the Hungarian/Italian/English dictionary of the COVID-19 pandemic (hereinafter referred to as the *Trilingual (HU, IT, EN) COVID-19 Dictionary*, TCD), I used a specialized *coronacorpus* extracted from the Web using Sketch Engine.³ To detect the related terms, I also analyze the Hungarian web corpora of news articles (online press) obtained from crawling a list of RSS feeds (Timestamped JSI web corpus).⁴ It is already highly evident that the vocabulary used in these articles (in the printed versions as well as in online press and media) is rather different with respect to the past. In fact, it is possible to note a frequency increase (for a short period, such as from March to the end of May 2020, or for a longer period, such as from March to the end of 2020) for certain word forms that are to some extent related to the all-encompassing COVID-19 pandemic. It is also possible to discover word forms that, before the outbreak of the pandemic, have never been seen in everyday

1 <https://www.lexonomy.eu/p8mwpck/> (last access: 10 June 2022).

2 The COVID-19 inspired neologisms or ‘coronacoines’ are sometimes referred to also as ‘coroneologism’, e.g., in papers written by Roig-Marín (2020). Previously, the term ‘coroneologism’ appeared in newspaper articles (e.g. Coroneologisms are going viral. In: *Economic Times*. April 9, 2020).

3 <https://www.sketchengine.eu/> (last access: 10 June 2022).

4 <https://www.sketchengine.eu/jozef-stefan-institute-newsfeed-corpust/> (last access: 10 June 2022).

Judit Papp, Department of Literary, Linguistic, and Comparative Studies, Palazzo Santa Maria Porta Coeli, University of Naples L’Orientale, Italy. Via Duomo 219 – 80138, Naples, Italy.
e-mail: jpapp@unior.it

Hungarian language. Many terms that usually belong to the medical and scientific fields (*epidemiology, virology, serology, etc.*) are being used in everyday language (in the press but also in informal contexts).

For the domain-specific terminology extraction, I used the Oneclick Dictionary function of Sketch Engine and created the first drafts of TCD.

From the dictionary drafts, I extracted the headwords related to the pandemic and included them in the TCD. I customized the structure and formatting of the dictionary in Lexonomy as well as configured the connection with my Sketch Engine account to have the possibility to extract and pull example sentences from Sketch Engine.

Finally, I completed the entries with the Italian and English equivalents and the corresponding examples taken from the Web as well as from the corresponding Timestamped JSI web corpora.

2 Field of study

Studies and research dedicated to the methodical lexicographic treatment of Hungarian terms related to the COVID-19 pandemic are still rather uncommon. The existing glossaries or dictionaries are usually monolingual (Hungarian) or bilingual (Hungarian-English). Except for my own *Trilingual (HU, IT, EN) COVID-19 Dictionary* (TCD), there is no other existing Hungarian-Italian dictionary on COVID-19.

Following the first wave of the pandemic, in 2020, a dictionary on the lexicon of COVID-19 was published in Hungary by Ágnes Veszelszki, the *Karanténszótár*, which collects 400 neologisms (words and expressions) that have appeared in the Hungarian language between January and July 3, 2020. Each lemma is accompanied by an explanation and examples taken from real texts. Besides the most commonly used words and expressions, Veszelszki has also included rather rare forms as well as *hapax legomena* in her dictionary. The dictionary, accompanied by a short essay, is an authentic and important testimony of the period under review, as it offers users a detailed view of the Hungarian linguistic aspects of the COVID-19 pandemic. In addition, it constitutes a valuable source for further linguistic reflections on the formation of neologisms in the Hungarian context. The essay is also interesting for the lexicon used by Veszelszki, as it is profoundly influenced by the pandemic and has its own related neologisms, namely *karanténszótár* ‘dictionary on quarantine’, *karanténszókinccs* ‘lexicon on quarantine’, *kórlenyomat* ‘imprint of the disease’, and *karanténkor* ‘period of quarantine’. These neologisms appear either in the title, introduction, or essay of the publication, but the author does not lemmatize nor define them among the items collected.

Another noteworthy publication on the challenges posed by COVID-19 and the various responses to the pandemic is *Globális kihívás – lokális válaszok* (Global challenge – local responses) edited by László Kovács (2020), which includes a section

dedicated to the articles that reflect on the new phenomena in the Hungarian language (Balázs; Domonkosi-Ludányi; Kegyes- Lanzmaier-Ugri and Lénárt).

With the creation of the TCD, my aim has been to fill this lexicographic gap primarily concerning the Hungarian-Italian language pair and to organize this content in a free online tool (a rich database) that is easy to search and useful for linguists and translators. The dictionary created with Lexonomy makes it possible to store, maintain, and update data in an organized manner. The third language is English, as the comparison with it is inevitable. On the one hand, that is due to the enormous quantity of news produced and conveyed, with extraordinary speed, by international agencies, a phenomenon exerts a considerable influence on other languages and, on the other hand, because English is the language of the international scientific community including, therefore, international medical research. The papers, findings, and results of scientists' experiments relating to COVID-19 are published in English, which means that English plays an important role in the creation of neologisms. In both Hungarian and Italian, we record a certain number of loans, calques, and adaptations, but we also have to deal with the needs of ordinary people and the creative abilities of individual languages.

3 Methodology

3.1 Corpus Selection

The COVID-19 Open Research Dataset (CORD-19) that is available on Sketch Engine consists of a collection of texts in English. As of November 2021, I still cannot find any specific Hungarian or Italian COVID-19 related corpus or a Hungarian-Italian COVID-19 related dictionary.

Therefore, for creating TCD, I decided to build my own COVID-19 related Hungarian corpus (in Hungarian *koronakorpusz*) using Sketch Engine and starting from the Web, as it represents an enormous resource ('web as corpus', cf. Kilgarriff 2001, Kilgarriff-Greiffenstette 2003). The Hungarian coroneologisms and words related to the COVID-19 pandemic are detected in this specific corpus that is built using all three options of Sketch Engine that make it possible to make the corpus larger:

- content downloaded by providing some typical words defining the topic (seed words),
- content downloaded by providing a list of URLs for download,
- content downloaded by downloading a complete website.

Content downloaded by providing the typical words that define the topic (seed words)

(“Find texts on the Web” option, input type: “Web search”)

As a result of this option, Sketch Engine extracted a series of web pages and documents. In “Web search”, I input groups of words and phrases (maximum 20) to enable defining the topic of the new corpus. With the pandemic’s progression and the succession of the different phases and waves, among others, I used seed words such as: *Astrazeneca*, *átoltottság* ‘vaccination coverage rate’, *COVID*, *COVID-19*, *COVID-igazolvány* ‘COVID certificate’, *delta*, *deltavírus* ‘delta variant’, *digitális* ‘digital’, *fertőtlenítés* ‘disinfection’, *fertőzés* ‘contagion, infection’, *fertőzött* ‘infected’, *görbe* ‘curve’, *harmadik* ‘third’, *hullám* ‘wave’, *immunitás* ‘immunity’, *járvány* ‘epidemic’, *karantén* ‘quarantine’, *koronavírus* ‘coronavirus’, *Moderna*, *mutáció* ‘mutation’, *oltás* ‘vaccination’, *oltásellenes* ‘anti-vax’, *oltáspárti* ‘pro-vax’, *oltópont* ‘vaccination point’, *Pfizer*, *Sputnik*, *szájmaszk* ‘mask’, *távolságtartás* ‘social distancing’, *tömeges* ‘massive’, *vakcina* ‘vaccine’, *vakcinabeszerezés* ‘vaccine procurement’, *védettség* ‘immunity’, *védőoltás* ‘vaccine’, *vírusvariáns* ‘virus variant’, etc.

An advantage of Sketch Engine is making it possible to run the corpus building tool many times to make it increasingly larger. It is also possible to repeat the search with the same seed words multiple times but also with different seeds, as well as to have multiword expressions using the quotes or proper names of different kinds. These words (seeds) are randomly selected and groups of three are sent to the Bing search engine. The Web pages that Bing returns are downloaded and processed into a corpus.

Content downloaded by providing a list of URLs that should be downloaded

I have also collected Hungarian language data from relevant URLs (e.g. blogs, forums, general websites on COVID-19, etc.). The main criterion for inclusion in the corpus is texts dealing with topics related to the pandemic.

Content downloaded by downloading a complete website

In particular, I have downloaded a few websites (July 12, 2021) containing useful information on the topic.⁵

⁵ The downloaded websites are: (i) <https://koronavirus.gov.hu/>, the official governmental portal in Hungarian on COVID-19 created together with the Operational Force responsible for the Prevention of the COVID-19 pandemic (*Koronavírus-járvány Elleni Védekezésért Felelős Operatív Törzs*) on January 31, 2020. The Ministry of Interior of Hungary is responsible for the operation of the portal and the Prime Minister’s Office is responsible for editing the content; (ii) <https://www.covid1001.hu/>: in the middle of March 2020, a group of medical translators (specialists, biologists, pharmacists, epidemiologists, language specialists) decided to combat misinformation by translating and publishing reputable articles; (iii) <https://semmelweis.hu/koronavirus/>, Semmelweis University’s website on the novel coronavirus, which is constantly updated with the latest news, information, communications, instructions, and actions concerning university citizens; (iv) <https://www.elte.hu/content/koronavirussal-kapcsolatos-tajekoztatok-cikkek.c2c.316>: Eötvös Lóránd University’s website that contains information on COVID-19 (updated: June 30, 2021);

With the help of its corpus building tools, Sketch Engine processed many Web pages and documents and built the Hungarian ‘coronacorpus’ (about 4 million words).

For the domain-specific terminology extraction, I used the Oneclick Dictionary function of Sketch Engine and created the first drafts of TCD. The Oneclick Dictionary is useful in automating the exchange of lexicographic data between the selected Sketch Engine corpus and a Lexonomy dictionary (Měchura 2017), even if post-editing is required. Besides my own specialized corpus, I analyze the following Hungarian web corpora of a news articles obtained from crawling a list of RSS feeds: *Timestamped JSI web corpus 2014–2020 Hungarian* and *Timestamped JSI web corpus 2021–01 Hungarian*. “The JSI Newsfeed corpus is a new family of Web corpora created from the JSI newsfeed of Jozef Stefan Institute, Slovenia [. . .]. JSI newsfeed is a clean, continuous, real-time aggregated stream of semantically enriched news articles from RSS-enabled sites across the world.” (Bušta et al. 2017). The corpora are tagged by TreeTagger v2.

Concerning *Timestamped JSI web corpus 2014–2020 Hungarian*, I have also created a sub-corpus that contains only articles from 2020, including 309,663,951 tokens and 256,156,393 words. The *Timestamped JSI web corpus 2021–01 Hungarian* contains 113,132 documents, including 34,378,246 tokens, 28,376,390 words, 1,624,519 sentences, and 699,713 paragraphs. Although these corpora are obviously not exhaustive, given these figures and the wide coverage of Hungarian language sources, I conclude that the size of the corpora can be suitable for analyzing the phenomena and trends in the Hungarian online press.

While my *coronacorpus* is useful in detecting the Hungarian *coroneologisms* (accepted by the speech community) and occasionalisms (or ‘nonce words’ coined for a particular occasion, e.g. *aranymaszk* ‘gold mask’) used not only in newspaper articles and standard Hungarian texts (everyday, neutral, unmarked) but also on other websites (government, homepages, school/university, etc.), blogs, and social networks (Facebook, Instagram, etc.). In this three-way, colloquial language (slang, informal, familiar) and formal language (scientific, specialized, academic, literary) will also be represented in TCD.

The Hungarian Timestamped JSI web corpora are an outstanding tool to detect the behavior of the words or single word forms. ‘Trends’, in fact, is a feature of Sketch Engine “for detecting words that undergo changes in the frequency of use in time (diachronic analysis). Trends identify words whose use increases or decreases in time.”⁶

Alongside this feature, ‘Concordance’ is useful, mostly the ‘Distribution of hits in the corpus’ function provides highly informative results. The ‘Word Sketch’ option, a

(v) <https://europa2000.hu/covid-19/>. The COVID-19 section of the website operated by the Europa 2000 Secondary School (Budapest), which is a secondary grammar school and vocational institution maintained by a foundation; and (vi) <https://www.pfizer.hu/>, the Hungarian version of the institutional site of Pfizer, one of the world’s premier innovative biopharmaceutical companies.

6 <https://www.sketchengine.eu/guide/trends/#toggle-id-6>.

one-page summary of the word's grammatical and collocational behavior, is another helpful feature (active in my own corpus, not available in the Timestamped corpora).⁷

A good example for illustrating the 'Trend' feature is the Hungarian coroneologism *nyunyóka*. The explosive growth of its frequency is strictly related to the pandemic. A *nyunyóka* can be anything that is safe for a baby or toddler to have at sleep time. It is a sort of comfort or transitional item – a blanket or stuffed animal or other comfort object of affection that a baby or toddler brings to bed, and that provides comfort and soothing. Previously, the term *nyunyóka* was uncommon and was used only in baby talk, and then, due to the massive media impact of Chief Medical Officer Cecília Müller's discourse during a press conference of the Operational Force, concerning personal hygiene habits to teach kids and the necessity to wash comfort objects frequently, this neologism entered the common language and became widely known and used. The number of hits found in the corpus is 166, for a lemma present only since May 13, 2020.⁸ Müller shared these tips instead of the daily COVID numbers, mortality and recovery rates, current active cases, recoveries, etc. that people were actually expecting. The results of the search query using Google now list 46,500 pages (as of December 16, 2021).

Besides the Timestamped JSI web corpora, the Web is a valuable corpus to find coroneologisms and forms belonging to Hungarian slang or to the colloquial register. The latter forms are usually not represented in current corpora typically based on news articles, which is why the creation of the Hungarian *coronacorpora* is important for this research.

From these drafts, I extracted the headwords related to the pandemic and included them in TCD. I customized the structure and formatting of the dictionary in Lexonomy and configured the connection with my Sketch Engine account so that there is an option to extract and pull example sentences from it. This option allows you to detect, select, and pull not only definitions and descriptions of the Hungarian coroneologisms (new words, new meanings of existing words, and new multiword units) into Lexonomy, but also collocates and collocations, etc. While building the dictionary, particular attention is paid to neologisms related to aspects regarding the outbreak of the pandemic, lockdowns, curfews, quarantines, social distancing, good hygiene practices, epidemiological curves, smart working, distance learning, first

⁷ However, in 'Show visualization', it would be great if the image could be editable by the user.

⁸ From Müller's discourse (https://index.indavideo.hu/video/Csenjuk_el_a_gyermek_nyunyokajat): "Tudjuk jól, hogy a piciknél van valamiféle ragaszkodás: itt nemcsak a cumikra gondolok, hanem kis pelenkára, vagy nyunyókára, amit ő otthonról hoz és nagyon szereti. Próbáljuk meg ezeket otthon gyakran tisztítani, elcsenni ameddig alszik a gyermek és ezeket kimosni és vasalással még egy hőkezelésnek alávetni, ami szintén fertőtlenítő hatású." (We know very well that little babies have some kind of attachment. Here, I am thinking not only of the pacifiers, but also of the little diaper or any comfort object he or she brings from home (to the nursery) and loves it very much. Let's try to clean them frequently at home, sneak them away from the child while they are asleep and wash them and subject them to heat treatment with ironing, which also has a disinfectant effect.)

wave, second wave, burden on healthcare systems, vaccines, and vaccine efficacy, third wave, fourth wave, variants, green pass, and the EU digital COVID-19 certificate. These aspects are where the largest part of new words came into existence. Common Hungarian terms that are important for understanding the COVID-19 pandemic are also included in the dictionary.

As far as the Italian and English equivalents are concerned, I proceed with interrogating the available Timestamped JSI web corpora for Italian and English and the above-mentioned COVID-19 Open Research Dataset (CORD-19): *Timestamped JSI web corpus 2014–2020 Italian*, *Timestamped JSI web corpus 2021–01 Italian*, *Timestamped JSI web corpus 2014–2020 English*, *Timestamped JSI web corpus 2021–01 English* and the *COVID-19 Open Research Dataset (CORD-19)*.

3.2 Terminological extraction

To be able to extract more and more COVID-19-related terms with my COVID-19-related Hungarian corpus and the Timestamped JSI web corpora, I used the ‘Keywords’ function (terminology extraction) that is available on Sketch Engine, downloaded and analyzed the ‘Wordlist’ (frequency list), and used the ‘Concordance’ function. In particular, in ‘Wordlist’ (BASIC tab), I searched for certain strings, such as *COVID*, *korona*, *karan* (from *karantén* ‘quarantine’), *járvány* (‘pandemic’), *vírus*, *fert* (from *fertőz* ‘infect’), *beteg* (‘ill’), *véd* (‘protect’), *olt* (‘to vaccinate’), *vakcina* (‘vaccine’), and *immun* to get the productivity of the corresponding lemmas.

3.3 Draft dictionary and formatting

Again, with my own COVID-19 related Hungarian corpus, I used the One-Click dictionary (automatic dictionary drafting) function of Sketch Engine to create my draft dictionary for Hungarian. The result was useful. From the draft, I extracted many headwords related to the pandemic. However, after a while, I learned how to use Lexonomy as well as how to configure and customize the dictionary structure and formatting, then I preferred to create a new, empty dictionary using the ‘Create a dictionary’ option and insert data manually, one by one. This method is time-consuming, but the content is more professional. I have also configured the connection with my Sketch Engine account to connect TCD with one of the corpora and implement the information available for the single terms or expressions.

3.4 The structure of TCD

Headwords consist not only of single words, but they also include particularly frequent or relevant multiword expressions (MWEs). Less frequent MWEs are presented as collocations of the headword or among the examples.

TCD is linked to the original corpus in Sketch Engine, and it is possible to detect, select, extract, and automatically pull definitions, examples of usage, collocations, and thesaurus items of the Hungarian coroneologisms from my corpus into Lexonomy.

Common Hungarian terms that are important for understanding the COVID-19 pandemic are also included in the dictionary, such as *járvány* ‘epidemic’, *vírus* ‘virus’, *fertőzött* ‘infected’, etc.

The Italian and English equivalents are added manually along with useful examples taken from texts on the Web.

4 First results

The use of several pre-existing occasional words and expressions has increased significantly during the COVID-19 period, while neologisms linked to the pandemic were coined with surprising speed (e.g. *covidiot*, *coronababy*, *zoom-kocsma* ‘virtual pub in Zoom’, *fotelviroológus* ‘armchair virologist’, etc.).

The lexical innovation resulting from the explosion of the pandemic is incomparable, as terms inspired and/or linked to COVID-19 entered the large-scale public consciousness. Faced with the new reality, the neologisms represent a functional tool to discuss all of the different phenomena related to the pandemic: the impact that the pandemic and the crisis have on our lives, society, and economy, the experiences following restrictive lockdown measures, and the many themes related to distance learning or vaccines. They are also useful for expressing our feelings or making light of our experiences.

While, on the one hand, words and expressions that have dominated the pandemic-related discourse since the outbreak of the pandemic have an informative function, on the other hand, in a certain sense they also allow us to gain mutual understanding, to protect each other, to share warnings, to comment on events, to express and share with others anxieties, fears, worries, anger or exasperation. With their help, we can also make jokes, laugh, or make fun of this shared lexicon or even rid ourselves of fears. For this purpose, my COVID-19 related Hungarian corpus can be useful, as it contains language data also from social media sites, forums, and blogs. In such texts we register some of the most common Hungarian COVID-19-related words with negative connotations, such as *COVIDszopás* ‘annoyance/unpleasant situation due to COVID’ (*szopás* means ‘sucking’), *COVID-tálibok* ‘COVID-Talibans’ (also

karanténtálibok ‘quarantine Talibans’), *COVID-fasizmus* ‘COVID fascism’, *COVID-faszszág* or *kovidfaszszág* ‘COVID bullshit’, *COVIDgeci* ‘unpleasant situation due to COVID’, etc.

The outbreak of the emergency in Hungary and its *maszk*-related “viral” lexicon

The second COVID-19 wave in Hungary began already in August 2020, and the infections were increasing exponentially. During this second wave, the emphasis was on the importance of wearing face masks (*maszkviselés*). Before the pandemic, the word *maszk* ‘mask’ is present in the corpus in minimal proportions and with a different meaning (‘a covering for the face that hides the person wearing it’), such as in the following example: “The robbers wore masks to hide their identities.” After the outbreak of the pandemic, due to the mandatory wearing of face masks, the use of *maszk* became widespread and its productivity exploded. With respect to *maszk* ‘mask’, *szájmaszk* ‘mouth mask’, and *arcmaszk* ‘face mask’, the word *arcpajzs* ‘face shield’ had no success and it did not spread in the common language (it registered only 478 hits in 2020). Among these words, the most frequent is *maszk* ‘mask’ (inflected forms included) with its 576,731 hits in the corpus. The following paragraphs show the high productivity of the term *maszk* and the frequency of the corresponding mask-related neologisms.

A group of these neologisms denotes different types of masks (*maszktípus* ‘mask type’ with 44 hits) according to (i) the area covered by the mask, (ii) its functions, (iii) the materials it consists of, etc.:

- (i) *szájmaszk* ‘mouth mask’ (9,774);⁹ *arcmaszk* ‘face mask’ (3,110); *orr-szájmaszk* ‘nose-mouth mask’ (20); *orrmask* ‘nose mask’ (5).
- (ii) *védőmaszk* ‘protective mask’ (3,605); *oxigénmaszk* ‘oxygen mask’ (152); *sebészmaszk* ‘surgical mask’ (100); *légzőmaszk* ‘respirator, breath mask’ (19); *vírusmaszk* ‘virus mask’ (10); *műtősmask* ‘surgical mask’ (5).
- (iii) *textilmaszk* ‘textile mask’ (413); *szövetmaszk* ‘mask made with fabric’ (142); *FFP-maszk* ‘FFP mask’ (23); *pamutmaszk* ‘cotton mask’ (19); *pleximaszk* ‘plexiglass mask’ (18); *vásznonmaszk* ‘canvas mask’ (17); *FFP3-maszk* ‘FFP3 mask’ (5).
- (iv) *csodamaszk* ‘miraculous mask’ (5).

A second group concerns the act of wearing the mask over the nose, mouth, and chin. The noun *maszkviselés* ‘wearing a mask’ (9,913) refers to the act of wearing a mask such as in *A maszkviselés kötelező marad kültéren is*. ‘Wearing (face) masks remains mandatory also outdoors’. In the corpus, there are different synonyms and

⁹ The numbers between brackets indicate the number of hits registered in 2020.

variants: *maszkhasználat* ‘usage of masks’ (1,805) and *szájmaszkhasználat* ‘usage of mouth masks’ (7); *maszkhordás* ‘wearing of masks’ (338); *maszkviselet* ‘wearing of masks’ (286) and *szájmaszkviselet* ‘wearing of mouth masks’ (5); *szájmaszkviselés* ‘wearing of mouth masks’ (71); *arcmaszkviselés* ‘wearing of face masks’ (8). To these abstract nouns we can add the derivational suffix *-i* to create adjectives: *maszkviselési* (1246) and *maszkviselési-* (4) ‘mask wearing’. An example is *maszkviselési szabályok* ‘mask wearing rules’. Other variants are *maszkhasználati* ‘mask usage’ (61), such as in *maszkhasználati szabályok* ‘mask usage rules’ or *maszkhordási* ‘mask wearing’ (42), cf. *maszkhordási fegyelem* ‘mask wearing discipline’; *szájmaszkviselési* ‘mouth mask wearing’ (11). More complex neologisms are the abstract noun *maszkviselés-ellenesség* ‘anti-mask wearing’ (2) and the adjective *maszkviseléses* ‘mask wearing’ (2), such as in *maszkviseléses élet* ‘mask wearing life’.

A person wearing a mask is *maszkos* ‘masked’ (1,272); *szájmaszkos* ‘masked with mouth mask’ (137); *védőmaszkos* ‘masked with protective mask’ (35); *arcmaszkos* ‘face masked’ (22); *maszkos-kesztyűs* ‘masked and gloved’ (12).

The Hungarian word *maszkviseelő* (34, present participle) can be used as an adjective or as a noun, cf. *Jómagam maszkviseelő állampolgár vagyok* ‘I am a mask wearing citizen’ or *Tudatos maszkviseelő vagyok* ‘I am a conscious mask wearer’.

The person who is not wearing a mask is *maszktalan* (*adj*) ‘without a mask’ (20), where *-talan* is a privative suffix or *maszknemviseelő* ‘person not wearing a mask’ (4). It is possible to add to the adjective *maszktalan* another derivational suffix to create the corresponding abstract noun *maszktalanság* ‘the condition of wearing no mask’ (2). The two compounds *maszknélküliség* (4) and *maszk-nem-viselés* ‘non-mask-wearing’ (1) have the same meaning. The last word can function also as a base for another adjective, *maszknemviselési* (*adj*) ‘not wearing masks’ (4), cf. *maszknemviselési vita* ‘debate around not wearing masks’.

An adverbial derivational suffix may be added to the adjective *maszkos* ‘masked’ as well: *maszkosan* ‘in mask; wearing a mask’ (21). E.g.: *Az viszont határozottan jó, hogy a diákok maszkosan nem tudnak cigarettázni!* ‘On the other hand, it is definitely good that students cannot smoke when wearing masks!’

In Hungarian, a person (or a group) that does not agree with wearing masks and spreads and encourages opinions against it is defined as being *maszkellenes* (180, *n, adj*); *maszkellenző* (3) or *maszktagadó* (92) ‘anti-mask’, e.g. *Magyarországon is vannak komoly maszkellenes csoportok* ‘There are also serious anti-mask groups in Hungary’; *Nem véletlen a rengeteg vírusszeptikus és maszkellenes, ám ezek nem a megfelelő reakciók egy ilyen válság idején.* ‘It’s no coincidence that there are plenty of viral skeptics and anti-masks, but these are not the right reactions in a time of such a crisis.’ Among the neologisms, there is also *maszkszeptikus* ‘mask skeptical’ (7) and *maszkhasználat-ellenes* ‘that does not agree with using masks’ (6). A particularly complex neologism is *maszkellenes-vírustagadó-konteós* ‘anti-mask, virus denier, conspiracy theorist’ (2).

With its 49 occurrences, the adjective *maszkmentes* ‘mask-free’ is also rather frequent in the corpus, e.g. *A vírustagadók egy része péntekre maszkmentes napot hirdetett*. ‘Some of the virus deniers declared a mask-free day for Friday.’ The corresponding abstract noun *maszkmentesség* ‘the state of being mask-free’ (5) is rather rare.

On the contrary, a person who agrees with wearing masks is *maszkrajongó* ‘fan of masks’ (6); *maszkipárti* ‘pro-mask’ (21) or *maszkhívő* (3), e.g. *Heves összetűzések voltak országszerte a maszkipárti és maszkellenes tábor között, gyakran az üzletek előtt*. ‘There were fierce clashes across the country between the pro-mask and anti-mask groups, often in front of shops.’

In the corpus we can also find the abstract nouns *maszktagadás* ‘negation of masks’ (15) and *maszkellenesség* ‘the condition of being anti-mask’ (16), and also *maszkvita* ‘mask debate’ (16) and *maszkháború* ‘mask war’ (8); the adjectives *maszk-elutasítási* ‘mask refusing’ (e.g. *a maszk-elutasítási hajlandóság az életkorral csökken* ‘the propensity for mask refusal decreases with age’); *maszkelutasító* ‘mask refusing’; *maszktagadó* (*adj*) ‘mask-denier’; *maszktalanítva* ‘unmasked; a person whose mask was removed’.

Maszk- ‘mask’ is present 130 times in wider expressions (with omissions), e.g. *Emellett a hétvégén újra kötelező a maszk- és kesztyűviselés*. ‘In addition, wearing a mask and gloves is mandatory again over the weekend.’

The compound noun *maszkgyártás* ‘production of mask’ is present in the 2020 Timestamped corpus 123 times, while the related *maszkgyártó* ‘producer of masks’ is recorded 120 times and *maszkgyár* ‘mask factory’ 7 times. It is also possible to find hits for *maszkgyáros* ‘mask manufacturer’ (4); *maszkgyáras* ‘manufacturer of masks’ (2), and *maszkgyár-látogatás* ‘mask factory visit’ (2). In the corpus, there are also a few *hapax legomenon* (with only 1 hit): *maszkgyári* (*adj*) ‘mask factory’ (e.g. *Trump még a maszkgyári programjára sem vett fel maszkot* ‘Trump did not even wear a mask during his mask factory visiting program’); *maszkgyártási* ‘mask manufacturing’ or *maszkgyártó-gép* ‘mask making machine’.

In addition, in the corpus, there are neologisms concerning:

- (i) the necessity to sew masks [*maszkvarrás* ‘mask sewing’ (95); *maszkvarró* ‘sewer of masks’ (32); *maszkvarrógép* ‘mask sewing machine’ (5); *szájmazskkészítő* ‘mouth mask maker’ (5)]; and
- (ii) the initial worldwide mask shortage and the necessity to provide masks to the population [*maszkhiány* ‘lack of mask’ (70); *maszkgigény* ‘mask demand’ (3); *maszkosztás* ‘distribution of masks’ (43); *maszkkészlet* ‘mask stock’ (18); *maszkbeszerzés* ‘purchase of masks’ (17); *maszkvásárlás* ‘mask purchase’ (15), *maszkpiac* ‘mask market’ (9); *maszkadomány* ‘mask donation’ (5); *maszkosztogatás* ‘mask distribution’ (5); *maszkdiplomácia* ‘mask diplomacy’ (59); *maszkalózkodás* ‘mask piracy’ (1); *maszkügy*, *maszkügy* ‘mask affair’ (18); *maszkszállítmány* ‘mask shipment’ (35); *maszkkiallítás* ‘mask exposition’ (5); *maszk-eladási* (*adj*) ‘mask sale’ (e.g. *Múlt hét végén drasztikusan megugrottak*

a hazai gyógyszertárakban az orvosi maszk-eladási számok ‘Last weekend, the numbers referring to the surgical mask sales in Hungarian pharmacies increased drastically’); *maszk-szállítás* ‘mask delivery’ (1); *maszkbiznisz* ‘mask business’ (6); *maszkeladás* ‘mask selling’ (3)].

In addition, the corpus contains 58 mask-related rare neologism and 53 *hapax legomena*. All these creations are included in TCD, even if this kind of data is usually left out of dictionaries. These *hapax* play an important role in the assessment of productivity and creativity of the Hungarian language. It is well known that the lifespan of a neologism is, from the moment of its first appearance, uncertain and difficult to predict: some of the neologisms seem destined to last, while others are not. In the long term, every forecast turns out to be uncertain and there is the risk of excluding neologisms destined for success. So, considering the neologisms containing *korona*- ‘corona-’, *COVID*-, *járvány*- ‘epidemic’, *vírus*- ‘virus’, *maszk*- ‘mask’, *karantén*- ‘quarantine’, and *oltás* ‘vaccine’ as constituents, I have decided to systematically collect all the new words encountered, without taking into account their actual use and the degree of their diffusion. In fact, there is a risk of including too many entries in TCD, but there is also the advantage of identifying with greater precision the paths of neological activity. Later in time, there will be the possibility to understand the reasons for the success of some of these new words and to discuss the predictable failure of many occasional and ephemeral creations. In any case, I consider it useful to record all of these neologisms, *hapax* included, even if I am aware that their neological status is objectively less strong and sustainable.

In total, in the corpus, there are more than 210 mask-related neologisms and only 9 are formed by ‘simple’ derivation, 1 is the lemma *maszk*, and the remaining parts are compounds and may be the result of multiple derivation. The neologisms *maszkné* ‘maskne’ and *maszkitisz* ‘maskitis’ are blend words (*maszk* + *akné*; *maszk* + *dermatitisz*) and they entered into the Hungarian language from English. *Aranymaszk* ‘gold mask’ is an occasionalism. It refers to the story of the businessman Shankar Kurhade who bought a customized gold mask. *Okosmaszk* ‘smart mask’ is formed following the examples of *okostelefon* ‘smart phone’, *okoseszköz* ‘smart device’, *okoszemüveg* ‘smart glasses’, *okosóra* ‘smartwatch’, etc.

All of the words are accompanied by useful child elements providing indications concerning frequency, such as ‘frequent’, ‘rare’, ‘hapax’, and frequency of use in time, such as a particular date (the point in time when a word started to be used), first, second, etc. wave or other information related to the trends of the words treated (unusual increase or decrease in use). For each entry, at least one translational equivalent will be provided in Italian and English. In Figure 1, the entry *maszkvita* ‘mask debate’ illustrates the microstructure of TCD.

maszkvita; n**Definition** pro és kontra a szájmazskokról**Word formation process** compounding**Frequency** rare**Trend** peaks in March and September/October 2020**Temporal use** from March 2020**Register** formal**Connotative effect****Subject field****Style** literal**Etymology****References**

— ♦ *Van olyan magyar település egyébként, ahol a polgármester döntötte el a **maszkvitát**: a koronavírus miatt szerdától csak maszkban mehetnek az emberek boltba és más üzletekbe a Fejér megyei Velencén.*

♦ *Emiatt hűzódott sokáig a **maszkvita** is, de most már szinte mindenki azt javasolja, elővigyázatosságból hordjunk maszkot ott, ahol másokkal találkozunk.*

Note *maszk + vita* “debate, discussion”

It. dibattito sulle mascherine; noun mwe**Etymology****References**

♦ *Il **dibattito sulle mascherine** è strettamente legato a un'altra questione che ha suscitato forti divisioni: in che modo il virus si sposta nell'aria e diffonde l'infezione?*

♦ *La variante Delta riaccende il **dibattito sulle mascherine** negli Usa*

Note**Eng. mask debate; noun mwe****Etymology****References**

♦ *Delta variant reignites US **mask debate***

♦ ***Mask debate** From School Boards to Courtrooms.*

Note

Figure 1: Entry *maszkvita* ‘mask debate’.

If, due to translational difficulties, no equivalent can be given, a descriptive/explanatory equivalent is added (cf. Figure 2), e.g. the word *nyunyóka* ‘comfort object’ has been in use since May 13, 2020, with peaks in May when it was used in the Chief Medical Officer Cecilia Müller’s discourse and in December 2020 when it was named the word of the year.

nyunyóka; n

Definition A nyunyóka is a lovey and it can be anything that is safe for a baby or toddler to have at sleep time. It is a sort of comfort or transitional item – blanket or stuffed animal or other comfort object of affection that a baby or toddler brings to bed, and that provides comfort and soothing.

Orthographic variants -

Word formation process derivation

Frequency Before May 13, 2020, the term nyunyóka was uncommon and was used only in baby talk.

Temporal use since May 13, 2020, as a result of the massive media impact of Chief Medical Officer Cecília Müller's discourse during the press conference of the Operational Force, concerning personal hygiene habits to teach kids and the necessity to wash comfort objects frequently, this neologism entered the common language and became widely known and used.

Trend Increasing. The results of the search query using Google now lists 46,500 pages (as of December 17, 2021).

Register baby talk > common language

Connotative effect term of endearment

Subject field

Style

Etymology Mind a nyanya, mind a nyunyó föltehetőleg dajka-, gyermeknyelvi, hangulatfestő és talán hangutánzó szó. Elképzelhető a nyúl, nyuszi szavak becézéséből fakad. Ezt erősíti, hogy a neten nyuszi-nyunyi szundikendő is rendelhető. <https://e-nyelvmagazin.hu/2021/01/10/nyunya-nyunyo/>

References Veszelszki 2020: 54–55.

— ♦ *Kerülni kell, hogy a gyerek otthonról játékot hozzon el: pelenka, „nyunyóka”, játék maradjon otthon, ám azokat otthon is tisztítsuk rendszeresen.*

Note Nyunyóka: Kisgyermek alvó játékszere, leginkább plüssfigura. Másnéven: alvóka, rongyi. Hangtani rokona a nyanya, a nagymama kedveskedő, az öregasszony gúnyos megnevezése.

It. doudou; n

Etymology

References

Note oggetto transizionale

Eng. lovey; n

Etymology

References

Note comfort object, transitional object

Figure 2: Entry *nyunyóka* 'comfort object'.

5 Conclusions

The high number of coroneologisms draw attention to the creativity and vitality of the Hungarian language in times of crisis, and the corpus analyses performed for the *Trilingual (HU, IT, EN) COVID-19 Dictionary* provide a clearer picture of the change in

the vocabulary during the COVID-19 pandemic and of the role and function of word formation processes that contributed to the creation of these neologisms. The analyses would suggest that the most frequently occurring word formation processes of the Hungarian neologisms related to the pandemic are compounding and derivation, but syntagms, blending, and semantic extension (changes in lexical meaning) are also used. [Furthermore, in Hungarian, new words may be productively created also by means of conversion, backformation, reduplication, clipping, loan words, and loan formations (e.g. calques), metaphor . . .] At the end of the pandemic, the analyses will also reveal to what extent the Hungarian language borrows coroneologisms from other languages.

Time passes, but the impact of the COVID-19 pandemic on the Hungarian language is still strong in 2021. It is true that many terms had their peaks during the first months of the crisis in 2020, but each phase and wave produce new topics, and terms, and generate considerable frequency increases in the use of certain forms. Therefore, the corpus-based dictionary will be a valuable tool to explore and analyze the coronalexicon in the Hungarian press and common language during this global emergency thanks to the microstructure of the entries that, where possible, includes information on frequency, trends, and temporal use. The entries contain a morphological analysis too, so that TCD provides data that will help to analyze the trends and patterns in the formation of new words and in their frequency of use in Hungarian.

Overall, this dictionary is useful for linguists and translators (e.g. suggesting more accurate translation equivalents for translating the coroneologisms from Hungarian to English or Italian) or for scholars in the digital humanities.

Bibliography

- Balázs, Géza (2020): A koronavírusról szóló beszéd (nyelv és folklór). In Kovács, László (ed.): *Globális kihívás – lokális válaszok. A koronavírus (Covid19) gazdasági és társadalmi összefüggései és hatásai*. Szombathely: Savaria University Press, 229–240.
- Bušta, Jan, et al. (2017): JSI Newsfeed Corpus. In: *The 9th International Corpus Linguistics Conference, July 25–28, 2017, Birmingham, GB: extended abstracts*. Birmingham: University of Birmingham, [<https://www.birmingham.ac.uk/Documents/college-artslaw/corpus/conference-archives/2017/general/paper382.pdf>; last access July 29, 2021].
- Coroneologisms are going viral. In: *Economic Times*. 9 April 2020. [<https://economictimes.india.com/blogs/et-editorials/coroneologisms-are-going-viral/> last access June 30, 2020].
- Dajka, Balázs (2020): Lecsapott a koronavírus az amúgy is nehéz időket élő kézilabdarúgócsapatra. In: *24.hu*. October 20, 2020. [<https://24.hu/sport/2020/10/20/kezilabda-koronavirus-karantenisofok-mtk/>; last access July 29, 2021].
- Domonkos, Ágnes/Ludányi Zsófia (2020): Társas távolságtartás és nyelvi közeledés – e-mailezési gyakorlatok a koronavírus idején. In: Kovács, László (ed.): *Globális kihívás – lokális válaszok. A koronavírus (Covid19) gazdasági és társadalmi összefüggései és hatásai*. Szombathely: Savaria University Press, 241–260.

- Elszabadult a koronavírus-járvány a világon: egy nap alatt 400 ezer fertőzött. In: *Portfolio.hu*. 19 October 2020. [https://www.portfolio.hu/gazdasag/20201019/elszabadult-a-koronavirus-jarvany-a-vilagon-egy-nap-alatt-400-ezer-fertozott-453448#; last access July 29, 2021].
- Hát ezért is olyan veszélyes, hogy berobbant Magyarországon a koronavírus. In: *Portfolio.hu*. 20 October 2020 [https://www.portfolio.hu/gazdasag/20201020/hat-ezert-is-olyan-veszelyes-hogy-berobbant-magyarorszagon-a-koronavirus-453454; last access July 29, 2021].
- Istók, Béla/Lőrincz, Gábor (2020): A virolingvisztika részterületei. In Simon, Szabolcs (ed.): *12th International Conference of J. Selye University. Language and Literacy Section. Conference Proceedings*. J. Selye University. Komárno. 83–92. [DOI: 10.36007/3761.2020.83; last access July 29, 2021].
- Kegyes, Erika/Lanzmaier-Ugri, Katharina (2020): A többnyelvű tájékoztatás kihívásai, stratégiái és eredményei a koronavírus idején Ausztriában. In Kovács, László (ed.): *Globális kihívás – lokális válaszok. A koronavírus (Covid19) gazdasági és társadalmi összefüggései és hatásai*. Szombathely: Savaria University Press, 261–289.
- Kilgarriff, Adam (2001): Web as corpus. In Rayson, Paul et al. (eds.): *Proceedings of the Corpus Linguistics 2001 Conference, Lancaster University, March 29–April 2, 2001*. Lancaster: UCREL, 342–344.
- Kilgarriff, Adam/Greiffenstette, Gregory (2003): Introduction to the Special Issue on the Web as Corpus. In: *Computational Linguistics* 3(29), 333–347. [DOI: 10.1162/089120103322711569; last access July 29, 2021].
- Kovács, László (ed.) (2020): *Globális kihívás – lokális válaszok. A koronavírus (Covid19) gazdasági és társadalmi összefüggései és hatásai*. Szombathely: Savaria University Press.
- Lénárt, István (2020): Karanténykedünk: a karanténdepitől a karanténszegig – avagy hogyan árulkodnak tevékenységeinkről a Koronavírus-járvány alatt született szóösszetételek. In Kovács, László (ed.): *Globális kihívás – lokális válaszok. A koronavírus (Covid19) gazdasági és társadalmi összefüggései és hatásai*. Szombathely: Savaria University Press, 291–302.
- Ludányi, Zsófia (2020): Helyesírási kérdések pandémia idején. In: *Amega: asztma és allergia színes tájékoztató magazin* 4(27), 33–35.
- Ludányi, Zsófia (2021): Eloltották az oltóanyagot. In: *Amega: asztma és allergia színes tájékoztató magazine* 2(28), 37–39.
- Ludányi, Zsófia (2021): Oltakozás és pándémia – avagy a virolingvisztika legújabb kérdései. In: *AMEGA: asztma és allergia színes tájékoztató magazin* 1(28), 43–45.
- Měchura, Michal Boleslav (2017): Introducing Lexonomy: an open-source dictionary writing and publishing system. In: Kosem, Iztok et al. (eds): *Electronic lexicography in the 21st century: Lexicography from Scratch. Proceedings of the eLex 2017 conference, September 19–21, 2017*. Leiden: Lexical Computing, 662–679.
- Papp, Judit (2021): il nostro lessico è diventato “virale”. Il vocabolario dell’emergenza sanitaria, economica e sociale ai tempi della pandemia di COVID-19. In: *Studia Universitatis Babeş-Bolyai – Philologia* 1(2021), 325–344.
- Roig-Marín, Amanda (2020): English-based coroneologisms: A Short Survey of Our Covid-19-Related Vocabulary. In: *English Today*, 1–3. DOI: 10.1017/s0266078420000255.
- Trampus, Mitja/Novak, Blaz (2012): The internals of an aggregated web news feed. In: *Proceedings of the Fifteenth International Information Science Conference IS SiKDD*. October 8–12, 2012, [Ljubljana, Slovenia], 431–434.
- Tyrrell, David Arthur John/Michael Fielder (2002): *Cold Wars: The Fight Against the Common Cold*. Oxford: Oxford University Press.
- Veszelszki, Ágnes (2020): *Karanténszótár*. Budapest: Interkulturális Kft.
- Virology. Coronaviruses. In: *Nature* 220 (November 16, 1968): 650. [https://doi.org/10.1038/220650b0; last access 29 July 2021].