

PREFACE

One of the aspects of a terminologist's work is to design, manage, and constantly update various terminological resources. In general, these are glossaries or terminological databases, where the information is systematically collected, following specific methodologies in the field of terminology. Currently, terminological activity is closely connected to digitalization, allowing terminological data to be exchanged between specialists in the field and reused in other contexts. The growing interest in what is now called corpus linguistics has been made possible by the advent of the computer. Combining the two efficiently highlights linguistic regularities in a language, based on computerized corpora.

As early as 1991 Sinclair defined a corpus as “a collection of naturally occurring language text, chosen to characterize a state or variety of a language.” Later on, he developed this definition into “A corpus is a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research.” (2005). The following studies in the field highlighted the importance of corpora design for language analysis. (Williams, 2006; Zeng, 2006, Quérin, 2013, Vezzani, 2019, 2022).

Over the years, specialists have developed different types of corpora that help in the study of general or specialized language (FRANTEXT or CEFLE, for example). In 2007, the European Union launched its own terminological database, IATE (<https://iate.europa.eu>), available in 23 languages. Its purpose is to enhance transparency in the legislative process and facilitate communication among EU citizens. This tool provides translators and terminologists with free and easy access to EU-related terminology (8.7 million terms). Compared to its predecessor databases (EURODICAUTOM, EUTERPE, and TIS), IATE is both interactive (the information in the database can be completed and updated by any EU translator) and interinstitutional (several EU institutions develop and maintain this database). For example, following the COVID-19 pandemic, terminologists from EU institutions have been interested in this area, leading to the addition of 270 multilingual entries related to the pandemic in IATE.

Regarding medical terminology resources, we mention:

- The TriMED database (<https://shiny.dei.unipd.it>), designed in three languages: Italian (436 terminology sheets), English (410 terminology sheets) and French (289 terminology sheets). The website states that the resource is intended for:
 - patients, to help find equivalents for technical medical terms;
 - translators, who can consult terminology sheets that provide information in both the source and target languages

- doctors, as a useful tool for interoperability in medical communication.

It is worth mentioning here that the TriMED project is the result of an interdisciplinary collaboration between Federica Vezzani (PhD candidate at the Department of Linguistic and Literary Studies of the University of Padua) and Professor Giorgio Maria Di Nunzio (member of the Information Management System group, Department of Computer Engineering of the University of Padua), who implemented this terminological database as a Shiny web application.

- The website <https://www.hetop.eu/hetop/rep/fr/terminologies/> which includes a list of 45 terminological-ontological resources, available in several languages, mainly English and French. Romanian is minimally represented, appearing more precisely in a single resource (<https://www.hetop.eu/hetop/rep/fr/CISP-2/>). This resource aims at an international classification in the field of primary care and includes a total of 745 concepts.
- For the Romanian language, the only notable resource is ROMBAC (The Romanian Balanced Annotated Corpus). This corpus includes texts from different fields (law, journalism, literary criticism), including medical domain. It was developed in 2012 by a team of the Romanian Academy Research Institute for Artificial Intelligence “Mihai Drăgănescu” (ICIA).

The poor representation of electronic Romanian language corpora, the rapid terminological evolution and lexical creativity during the COVID-19 health crises (March 2020-May 2023) are the main reasons that led us to initiate a project aimed at analyzing the intralinguistic and interlinguistic terminological evolution and developing a multilingual glossary (English, French, German, Romanian) that would bring together terms used in any context related to the pandemic.

The project was carried out by a team from the Logos Foreign Language Centre from the National University of Science and Technology POLITEHNICA Bucharest in collaboration with specialists in the field of medicine and computer science from the same university.

The innovative character and original elements proposed by the *PandemicTerm* project consisted of the following aspects:

- the multilingual corpus (English, French, German, Romanian) based mainly on about one hundred texts (about 1 million characters) that come from official documents by authorities;
- manual processing assisted by computer and automatic processing using terminological extraction software, through operations like segmentation and annotation with morphosyntactic labels;
- the standardization of the document collection, in terms of both graphics and content;

- the design, management, permanent updating of a terminology database (in Excel document and online) that targets the already mentioned field;
- the interdisciplinary character of the project will help mainly in the validation stage of the terms and in the establishment of the database from a technical point of view, by collaborating with specialists from the medical and computer science fields.

From the outset, we emphasized the interdisciplinary nature of the activities carried out within the project, with each of the three teams having clearly defined roles:

- For the researchers in linguistics, translation, and terminology, the main responsibility was to compile the corpus, to identify candidate terms, and analyze terms with varying degrees of specialization in English, French, German, and Romanian;
- For the researchers with expertise in the field of medical sciences, the primary responsibility was the validation of selected terms, which formed a micro-corpus of terms, classified according to the following stratification: general medical terms, medical terms common to several branches of the medical field, and medical terms with a narrow niche of specialization.
- The task of the computer science research team was the creation of the terminological database with several mandatory sections: entry number of the chosen term, domain, linguistic and etymological information, definition, references, contextual examples, and equivalent terms in other languages.

Our approach was therefore part of a textual approach based on a “methodological tripod” (Picton & Dury 2015) which uses “texts (corpus to be compared), tools and indices (whose terms are the entry points for the analysis) and interviews with experts (in order to co-construct an interpretation of the data)” (Picton, 2018:3).

This contextualized multilingual glossary represents the publication in book form of the terminological database extracted from the analyzed corpus. It is intended to be a useful tool for translation and terminological activities, medical professionals and students, or simply a small testimony to one of the most important health crises of the last century.

ENGLISH

A

Ageusia

Ageusia is a symptom reflecting various underlying pathological conditions. It is a rare condition that is characterized by complete loss of the taste function of the tongue.

<https://www.ncbi.nlm.nih.gov/>

*There are a variety of conditions that can lead to **ageusia**, such as damage to the nerve of taste sensation (lingual and glossopharyngeal nerve) in the anterior and posterior portion, dietary deficiencies, systemic conditions such as hypothyroidism, and diabetes mellitus, pernicious anemia, Sjogren syndrome, and Crohn disease.*

<https://www.ncbi.nlm.nih.gov/>

Anosmia

Loss or impairment of the sense of smell.

<https://www.merriam-webster.com/dictionary>

*In a survey of 180 non-hospitalized COVID–19 patients, over 50% reported having at least one persistent symptom (most frequently fatigue and **anosmia**) approximately 4 months after the onset of symptom.*

<https://www.ncbi.nlm.nih.gov/>

Antibody

Antibodies are proteins made by your body's immune system to help fight off infections, including those caused by viruses. Some antibodies in your body may protect you from getting those infections. Your immune system can also safely learn to make antibodies through vaccination. If antibodies give you this protection and how long this protection lasts can be different for each disease and each person. Antibodies are just one part of your immune response.

<https://www.fda.gov/>

*It is unknown if all people who have a SARS-CoV-2 infection will develop **antibodies** in their bodies in an amount that can be detected by a SARS-CoV-2 antibody test.*

<https://www.fda.gov/>

Antigen

Antigens are substances that cause the body to produce an immune response – they trigger the generation of antibodies. These tests use lab-made antibodies to search for antigens from the SARS-CoV-2 virus.

<https://www.umassmed.edu/>

*To run an **antigen** test, you first treat a sample with a liquid containing salt and soap that breaks apart cells and other particles. Then you apply this liquid to a test strip that has antibodies specific to SARS-CoV-2 painted on it in a thin line.*

<https://www.umassmed.edu/>

Antiretroviral therapy (ART)

Antiretroviral therapy (ART) is treatment of people infected with human immunodeficiency virus (HIV) using anti-HIV drugs. The standard treatment consists of a combination of drugs (often called "highly active antiretroviral therapy" or HAART) that suppress HIV replication. The combination of drugs is used in order to increase potency and reduce the likelihood of the virus developing resistance. ART reduces mortality and morbidity rates among HIV-infected people and improves their quality of life. The benefits of ART also include the prevention of HIV transmission by suppressing HIV replication in persons living with the virus.

<https://www.paho.org>

*"Use digital mapping and other innovative solutions developed for COVID-19 vaccination (e.g. electronic individual immunization registries) to estimate the number of beneficiaries from high priority-use groups for other programmes and services, such as antenatal care, **antiretroviral therapy (ART)** and screening services for noncommunicable diseases."*

<https://www.who.int>

Anti-vaxxer

A person who opposes the use of some or all vaccines, regulations mandating vaccination, or usually both.

<https://www.merriam-webster.com/>

Anti-vaxxers see COVID-19 as an opportunity to rapidly create widespread hesitancy for vaccines against COVID-19 and, indeed, for all vaccines.

<https://dictionary.cambridge.org/>

Asymptomatic contact

People who are asymptomatic show no signs or symptoms of an illness or disease, such as COVID-19, but can still transmit the virus that causes the disease to others. Anyone who has come into contact with someone who tests positive for COVID-19 must get a test themselves.

<https://www.medicalnewstoday.com>

*In reviewing articles obtained from a MEDLINE search, the author found that no trials have been specifically performed to evaluate whether treatment of **asymptomatic contacts** or family members improves outcomes.*

<https://www.aafp.org>

Attack rate

Attack rate is really the proportion of exposed people that become ill. [...] An attack rate is used when the occurrence of disease among a population at risk increases dramatically over a short period of time.

<https://epiville.ccnmtl.columbia.edu>

*The **attack rate** (AR) is the cumulative incidence of cholera since the start of the outbreak. As the AR is population-based, knowing the total number of people in the affected area is essential.*

<https://medicalguidelines.msf.org/>