

LINGUISTIC RESEARCH  
IN THE FIELDS OF CONTENT DEVELOPMENT  
AND DOCUMENTATION

L'Harmattan Hongrie



# COLLECTION KÁROLI

*Collection dirigée par Enikő Sepsí*

ISSN 2062-9850

Collection Károli

The Hungarian Language and Professional Culture:  
Hungarian Linguistic Research in the Fields of Content Development  
and Technical Documentation – Volume 7

Series Editors: Ágota Fóris and Andrea Bölcskei

**LINGUISTIC RESEARCH  
IN THE FIELDS OF CONTENT  
DEVELOPMENT AND  
DOCUMENTATION**



EDITED BY

ÁGOTA FÓRIS – ANDREA BÖLCSKEI

Károli Gáspár University of the Reformed Church in Hungary

L'Harmattan Publishing • Éditions L'Harmattan

Budapest • Paris

2021

Publishing Director: Géza Horváth, Ádám Gyenes, Xavier Pryn  
Series Editor: Enikő Sepsi

Károli Gáspár University of the Reformed Church in Hungary  
Kálvin tér 9.  
H-1091 Budapest, Hungary  
T: (+36-1) 455-9060  
L'Harmattan Kiadó  
Kossuth Lajos utca 14–16.  
H-1053 Budapest, Hungary  
L'Harmattan France  
5-7 rue de l'Ecole Polytechnique  
75005 Paris

© Authors, Editors, 2021  
© Károli Gáspár Református Egyetem, 2021  
© L'Harmattan Kiadó, 2021  
© Éditions L'Harmattan

The publication of the book was funded by the project titled *The Hungarian Language and Professional Culture: Hungarian Linguistic Research in the Fields of Content Development and Technical Documentation* by the Faculty of Humanities and Social Sciences of the Károli Gáspár University of the Reformed Church in Hungary.

Reviewers

János-Imre Heltai, Eszter Kárpáti, Kinga Klaudy, Eszter Sermann

Translator  
Eszter B. Papp

Proofreader  
Peter Sabath

The cover features an open access photo from Wikimedia Commons of the painting by Tivadar Csontváry Kosztka titled *Sun Looking back at Trau*

ISBN 978-2-343-25167-7

Volumes may be ordered, at a discount, from:

L'Harmattan Könyvesbolt  
1053 Budapest, Kossuth L. u. 14–16.  
tel.: +36-1-267-5979  
harmattan@harmattan.hu  
www.harmattan.hu  
Librairie internationale  
16 Rue des Écoles, 75005 Paris  
librairie.internationale@harmattan.fr  
www.amazon.fr

# CONTENTS



ÁGOTA FÓRIS: Introduction .....	7
---------------------------------	---

## **Part I Technical Communication**

ÁGOTA FÓRIS: Content Development and Documentation from the Perspective of Hungarian Linguistics .....	11
ESZTER B. PAPP: Professional Documentation in the 21st Century .....	41
ESZTER B. PAPP: Documentation 4.0 .....	59
ANDREA BÖLCSKEI: Practices in Technical Communication and the Latest Challenges .....	75
ANDREA BÖLCSKEI: What Do Technical Writers Do and What Do They Learn? .....	95
ÁGNES CZINKÓCZKI – ILDIKÓ FEHÉR: How is Technical Content Created? The Lifecycle of Technical Documentation .....	119

## **Part II Documentation, Terminology and Translation**

ÁGOTA FÓRIS – ANDREA FALUDI: The Role of Documentation and Document Management in Translation and Terminology .....	139
ÁGOTA FÓRIS – ANDREA FALUDI: Technical Writing and Documentation as Intralingual Technical Translation .....	157
MÁRTA KÓBOR: Features of Texts on Websites and Characteristics of their Translation .....	177
ANDREA FALUDI: Terminology Work in the Process of Technical Writing and Translation. Results from an Interview Survey.....	203
DÓRA-MÁRIA TAMÁS: Terminology and Documentation in Terminology Databases .....	227
ESZTER SERMANN: Concepts Related to the Term <i>Documentation</i> in Cercaterm, a Catalan Terminology Database .....	239

**Part III**  
**Technical Texts**

CSILLA-ILONA DÉR: Scientific Genres – Specialized Genres . . . . .	255
NÓRA CSONTOS: Means of Sharing Knowledge in Technical Texts. A Possible Interpretation of the Technical Text Types . . . . .	269
NÓRA CSONTOS: Technical Texts from the Perspective of Text Typology. Text Typology Analysis of <i>Instructions for Use</i> . . . . .	293
ÁGNES CZINKÓCZKI – ILDIKÓ FEHÉR: Types and Characteristics of Technical Texts . . . . .	313
RÉKA SÓLYOM: Semantic Features in the Terms of the Hungarian Version of an EU Regulation . . . . .	333
RÉKA SÓLYOM: Conceptual Metaphors and Metonymies in a Technical Text on Quality Management Systems . . . . .	353
CSILLA-ILONA DÉR: The Use of Discourse Markers in Architectural Technical Documentation . . . . .	375
Contributors . . . . .	387

CONCEPTS RELATED TO THE TERM *DOCUMENTATION*  
IN CERCATERM,  
A CATALAN TERMINOLOGY DATABASE<sup>1</sup>

—◀▶—  
ESZTER SERMANN

*Over the last few years, the Catalan language – as well as the attempts at independence in the autonomous community of Catalonia – has been a subject of international interest. During Francoist dictatorship (1939–1975), the use of Catalan was forbidden at every official level, but as a result of the democratic constitution (1978) it gained a co-official status. One of the practical results of the conscious language and terminology development in recent years is Cercaterm, the Catalan terminology database. After presenting several changes in the status of Catalan language and analyzing Cercaterm, this study deals with concepts related to the term documentation in the database, with special regard to the concepts belonging to domains such as library science and document management.*

*Keywords: documentation, document management, Catalan language, terminology development, terminology database*

PURPOSE AND SOURCE OF RESEARCH

The term *documentation* is associated with several concepts in the fields of technical documentation and document management. The aim of the present study is to investigate the content and form of the entries on the term *documentation* in the Catalan terminology database Cercaterm. The database, which is mainly used for language policy purposes, was created by the Catalan Terminology Centre, Termcat, in order to share the results of the organization's systematic terminology work. After presenting the changes in the status of the Catalan language and providing a detailed analysis of the Cercaterm database, the study deals with the concepts related to the term *documentation* in the database, with special regard to the concepts belonging to domains such as library science and document management.

<sup>1</sup> The bibliographical data of the first, Hungarian publication of this paper: Sermann, Eszter: *A dokumentáció terminushoz kapcsolódó fogalmak a Cercaterm katalán terminológiai adatbázisban*, in Fóris, Ágota – Bölcskei, Andrea (eds.): *Tartalomfejlesztés és dokumentáció. Nyelvészeti kutatások*, Budapest, KRE–L'Harmattan, 2021, 247–262.

## CHANGES IN THE STATUS OF THE CATALAN LANGUAGE

Catalan is spoken by around 10 million people in four European countries, mainly in the north-east of Spain (Catalonia, Valencia, the Balearic Islands and parts of Aragon), in Andorra, in some parts of France and in the city of Alghero on the Italian island of Sardinia. Catalan belongs to the Romance language family. Its closest relative is Occitan, and its narrower classification as a Gallo-Romance or Ibero-Romance language is still the subject of typological debate among historical linguists.<sup>2</sup>

In recent years, several changes took place in the status and official use of Catalan and Spanish in Catalonia, and the practical aspects of linguistic autonomy have been reinterpreted several times in parallel with the issue of Catalan independence. The situation of Catalan in Spain is specific: its use, like that of Basque and Gallego languages, was banned in all official areas during the Franco era, following the Spanish Civil War. This set back the development of these historic languages for decades, before they were given co-official (*cooficial*) status in some autonomous communities. The theory and practice of terminology in Spain is uneven: the Catalan, Basque and Gallego languages, in their respective autonomous communities, have more organized terminology work than the majority Spanish language.<sup>3</sup> This may be due to the fact that after many years of silence these languages were rediscovered, and efforts to standardize terminology were launched in an organized way. Scientific and technical progress required language development and terminology planning, which meant that terminology work was launched under central, autonomous provincial control.

In Spain after 1978, following the adoption of the constitution, autonomous communities also drafted their own statutes of autonomy, and these documents also dealt in detail with the issue of linguistic autonomy. The Catalan Statute of Autonomy of 1979 states that Catalan is the official language of Catalonia and also recognizes Spanish as an official language. Following the drafting of the Statute of Autonomy, the Catalan Language Normalization Act was passed in 1983, which stressed the importance of Catalan taking its rightful place in official spheres related to education and public media. The language campaigns launched in the 1980s achieved significant results in these key areas by involving the population and emphasizing social responsibility. A new linguistic policy emerged in 1998 and a new Statute of Autonomy was drafted in 2006.<sup>4</sup>

<sup>2</sup> On the typological issues of the Catalan language, see Berta, Tibor: A katalán az iberoromán nyelvek között, in Anderle, Ádám (ed.): *Kutatási közlemények III. A magyar katalán kapcsolatok ezer éve*, Szeged, Hispánia, 2001, 87–95.

<sup>3</sup> Cabré, M. Teresa: Terminology Today, in Somers, H. (ed.): *Terminology, LSP and Translation*, Amsterdam–Philadelphia, John Benjamins, 1996, 15–34.

<sup>4</sup> The details of the Catalan language autonomy are discussed in detail in Szente-Varga, Mónika: Katalán nyelvpolitika, *Mediterrán világ*, Vol. 21, 2012, 53–63.

The change in the status of Catalan after the Franco era also changed the role of the language in education. By the 1980s, Catalonia had become a developed industrial hub attracting immigrants, and by the mid-1980s half of the population was comprised of immigrants. While the 1979 Act on autonomous communities allowed the use of Catalan in Catalonia, some of the teachers who taught in public schools did not speak Catalan or could not teach in Catalan. The 1983 Law for Linguistic Normalization in Catalonia, therefore, also stipulated that Catalan should be the language generally used in public education, and that schools should organize remedial courses for children in public education to balance out differences in both Spanish and Catalan. As a result of the 1998 law and the language laws of 2006, more and more subjects were taught in Catalan, and it became the language of everyday life and of public education. In 2014, the Spanish parliament drafted a bill to demand the use of more Spanish in Catalonia, like in the rest of the country, but the bill was rejected several times and it still has not become law.<sup>5</sup>

#### LANGUAGE AS A TOOL OF CATALAN IDENTITY

The history of the autonomous Catalan community in the twentieth century is a direct consequence of the tensions of previous centuries. Catalonia produces 20% of Spain's GDP, but the central government returns much less than that to the autonomous community. The 2008 world economic crisis further deepened the divisions between Catalonia and Madrid.<sup>6</sup> Language has always had an important role in the development of Catalan identity, and pro-independence politicians and some part of the society still see language as the primary symbol of identity, making it a tool in the current independence movement. The Catalan language has always been a political issue as well because language is an expression of identity that is often used as a tool by politicians. The 2006 Statute states that Catalan is the primary language used in Catalonia in public administration, media and education, and that residents of Catalonia must speak both Catalan and Spanish. It also codifies the principle that no one should be discriminated against on the basis of language.<sup>7</sup>

Language is an important part of Catalan identity, and Berta<sup>8</sup> claims that today the intentions of those who argue for independence or for unity are also

<sup>5</sup> Németh, Tibor: Regionális autonómia és nyelvpolitika Katalóniában, *Új Pedagógiai Szemle*, 2018/5–6, 116–122.

<sup>6</sup> The history of Catalonia in the 20th century is discussed in detail in Lénárt, András: Katalónia a 20. században, *RUBICON*, Vol. 28, 2018/11, 54–63.

<sup>7</sup> Pujol Berché, Mercè: Política lingüística: lengua cultura e identidad, el ejemplo de Cataluña, *Amnis, Revue d'études des sociétés et cultures contemporaines Europe-Amérique*, 2013/12, <https://journals.openedition.org/amnis/2061> (accessed June 30, 2020).

<sup>8</sup> Berta, Tibor: Nyelv és identitás: a katalán esete, *Szépirodalmi Figyelő*, 2018/1, 45–54.

reflected in language, as those who argue for independence do so in Catalan and those who argue for unity speak in Spanish. The co-official language status has created a situation of asymmetry in some respects, as Catalan language policy seeks to provide public education in Catalan with the aim of “re-Catalanizing” Catalonia. But this is disadvantageous for the non-Catalan-speaking population because it creates tension. Déri and Fabula<sup>9</sup> also identify language as an important component of Catalan identity, listing its own history, culture and pride in these as other factors. This, the authors argue, is the reason for striving for independence, rather than a fear of losing relative prosperity and high living standards. The lack of independence makes Catalans feel that it is difficult to guarantee the survival of their language, and that their cultural achievements in music, fine arts, architecture and sport are being lost in the shadow of being Spanish. Another reason why the Catalans hold their language so dear is described by Nemes.<sup>10</sup> The language will remain alive if it is actively present at all levels of language use in society, and bilingualism is considered a temporary state. The author finds it possible that Catalans will eventually switch languages in everyday interaction with a world language. Since Spanish is more widely spoken, it might become dominant in everyday communication, in business and in law. Catalan might recede to family and school use, which might lead to a scenario in which Spanish displaces Catalan in democracy as well in the long-term.

CERCATERM, THE DATABASE HOSTING THE TERMINOGRAPHIC ENTRIES  
CREATED BY TERMCAT

The Cercaterm database was created by the Catalan terminology center, Termcat. Termcat was founded in 1985 following an initiative of the Catalan Government (*Generalitat de Catalunya*)<sup>11</sup> and the Institute of Catalan Studies (*Institut d'Estudis Catalans*).<sup>12</sup> According to the website of Termcat,<sup>13</sup> its mission is to develop Catalan terminology and to organize the terminology of various subject areas by continuously creating innovative resources and tools, in constant contact with professionals and users of terminology. The next section of this paper analyzes the database and the entries on the term *documentation*. First, a definition of database is provided. Terminology databases are defined as collections of electronically stored terminology data that are cre-

<sup>9</sup> Déri, Balázs – Fabula, Kálmán: Fejezetek a katalán kultúra történetéből, *Szépirodalmi Figyelő*, 2018/1, 26–44.

<sup>10</sup> Nemes, Krisztina: Függetlenség – állam – nyelv: Katalónia különös esete, *Létünk*, 2015/2, 117–128.

<sup>11</sup> <http://web.gencat.cat/ca/inici> (accessed June 30, 2020).

<sup>12</sup> <https://www.iec.cat/> (accessed June 30, 2020).

<sup>13</sup> <https://www.termcat.cat/ca> (accessed June 30, 2020).

ated following an onomasiological (concept-based) approach and contain terms from one or more disciplines in one or more languages, together with their associated definitions.<sup>14</sup>

Databases created according to terminology principles also play an important role in the life of the community from the point of view of language policy and terminology policy. The UNESCO *Guidelines for Terminology Policies* describes terminology policy as a set of written strategies that aim to coordinate terminology use of various groups of society. One of the tasks set out in the document is to examine the best way to publish data on terminology development.<sup>15</sup> When they outlined the tasks of Hungarian terminology strategy, Fóris and Bölcskei<sup>16</sup> stated that a terminology database which is the result of cooperation between experts and terminologists is an essential tool for knowledge management and an effective knowledge repository. This is indeed how Termcat performs its tasks, by creating and making publicly available databases of terminology.

In the present survey, a comprehensive, multi-criteria system is used to introduce Cercaterm, which examines publicly available, central terminology databases along four categorical elements.<sup>17</sup> This evaluation system was developed based on criteria for the analysis of print and online dictionaries, and it is structured around the following criteria: 1) background information on the database (overview of databases available in the domain, classification by trend, information on the database owner, the target group and purpose of use); 2) technical parameters (nature of the software, database accessibility); 3) content information (general characteristics of the database, detailed characteristics of its structure); 4) information on the use of the database (ease of use, data

<sup>14</sup> Sermann, Eszter – Tamás, Dóra: Hogyan definiálhatjuk a fordítói adatbázist? Egy olasz és egy spanyol fordítói terminológiai adatbázis vizsgálata, in Károly, Krisztina – Fóris, Ágota (eds.): *Nyelvek találkozása a fordításban. Doktori kutatások Klaudy Kinga tiszteletére*, Budapest, ELTE Eötvös, 2010, 101–116.

<sup>15</sup> UNESCO: *Guidelines for Terminology Policies. Formulating and implementing terminology policy in language communities*, 2005, <https://unesdoc.unesco.org/ark:/48223/pf0000140765> (accessed June 30, 2020); Hungarian translation: UNESCO (Hungarian translation): Terminológiai politikai irányelvek. A terminológiai politika kialakítása és megvalósítása különböző nyelvközösségekben (original publication by: Infoterm, translated by Somogyi, Zoltán), in Fóris, Ágota – Bölcskei, Andrea (eds.): *Terminológiai stratégiai kihívások a magyar nyelvterületen*, Budapest, L'Harmattan–OFFI, 2019, 167–218, <https://www.offi.hu/offi-akademia/kiadvanyok/terminologiai-strategiai-kihivasok-a-magyar-nyelvterületen> (accessed June 30, 2020).

<sup>16</sup> Fóris, Ágota – Bölcskei, Andrea: Ajánlások a magyar terminológiai stratégiához, in: Fóris, Ágota – Bölcskei, Andrea (eds.): *Terminológiai stratégiai kihívások magyar nyelvterületen*, Budapest, OFFI–L'Harmattan, 2019, 140–164, <https://www.offi.hu/offi-akademia/kiadvanyok/terminologiai-strategiai-kihivasok-a-magyar-nyelvterületen> (accessed June 30, 2020).

<sup>17</sup> Tamás, Dóra-Mária – Sermann, Eszter: Elemzési szempontrendszer terminológiai adatbázisokhoz, *Fordítástudomány*, Vol. 21, 2019/2, 46–62; Tamás, Dóra-Mária – Sermann, Eszter: Evaluation System for Online Terminology Databases, *Terminologija*, Vol. 26, 2020, 24–46.

updating and reliability, database innovation, social utility and professional relevance).

### **Background information on Cercaterm**

The website of Termcat, the organization that hosts Cercaterm, is available in Spanish, Catalan and English, and its objective is to promote the standardization process of the Catalan language, respecting multilingualism and diversity. The organization is run by the Board of Directors (*Consell de Direcció*), and terminology standardization (i.e. the selection of preferred terms from among the term candidates necessary to disseminate new knowledge in the various subject areas) is carried out by the Board of Supervisors (*Consell Supervisor*). Termcat collaborates with Spanish and international organizations committed to terminology, promotes multilingual terminology work with a view to promoting Catalan in the international arena, works closely with universities in the Catalan speaking area, liaises with professionals involved in terminology projects, and contributes to university training by providing traineeship opportunities for undergraduate and postgraduate students in terminology, humanities, translation and interpreting. The other terminology database produced by the organization is Neoloteca, which contains neologisms standardized by Termcat.

Cercaterm serves both a language policy purpose and a standardization role. Termcat's tasks are defined by the Catalan Government as follows: coordinating terminology activities for Catalan, developing terminology tools and making them available to the general public; planning and carrying out terminology research related to Catalan; developing terminology tools to enable the use of Catalan in economic, scientific and technical fields; developing methodological principles for the development of terminology tools; monitoring and standardizing Catalan terms; introducing and disseminating neologisms; and creating and regularly updating terminology databases. The databases produced by the organization widely disseminate standardized terms, thus serving as a primary reference for clarifying concepts and the use of terms.

The database is intended primarily as a useful tool for translators, interpreters and documentation producers. The question of whether Cercaterm is a prescriptive or descriptive database is answered in the Cercaterm Help menu item, in the description of resources and working methods. Descriptive databases provide users with information on terminological data, whereas prescriptive databases are normative. Cercaterm has both features: on the one hand, it contains the concepts and terms standardized by the Board of Supervisors, in which case the prescriptive character becomes more dominant; and on the other hand, it brings together the results of terminology research carried out by Termcat, and can therefore be considered a descriptive database.

### Technical parameters of *Cercaterm*

The documentation on *Cercaterm*'s user interface does not make it clear what software was used to create it, but it does say that the institution makes its terminology corpus available to researchers and doctoral students, and that the *GesTerm* terminology management software can be downloaded free of charge from its website, which allows the main tasks related to terminology work to be carried out: creating terminographic entries and recording information related to them. Their software, called *GdTweb*, is designed for terminologists and professionals who work remotely with *Termcat* and allows them to use a database management program linked to the Centre's terminology database online. It is therefore assumed that proprietary software was also used to develop the database.

The *Cercaterm* database is open to the public and accessible to anyone without registration, as it aims to disseminate the terminological data standardized by the institution as widely as possible. Users can register on the site, which serves as a user interface for the database, and once registered, they can contact *Termcat*'s customer service, make suggestions and send comments. There is no dedicated mobile phone application developed for the database, but the database interface is optimized for mobile phones and can be easily consulted in the phone's browser. On the website of the database, clicking on the *Terminologia Oberta* (Open terminology) tab opens a user search surface similar to *Cercaterm*, which allows users to download data originating from *Termcat*'s terminology research in different file formats (PDF, XML, HTML). An advantage of making it downloadable is that the resulting data can be imported into terminology management software and used effectively during translation.

### Information on the content of *Cercaterm*

In order to describe the general characteristics of the database, it is worth considering whether semasiological or onomasiological principles have been applied in its development. Looking at the user interface of *Cercaterm*, the concept-oriented approach of editing is evident from the fact that all entries are classified by subject and definitions are provided in at least one language, which makes *Cercaterm* an onomasiological database. The general characteristics of any database show whether it is simple, traditional or complex,<sup>18</sup> and in this sense *Cercaterm* can be considered a traditional database: it is concept-oriented, it observes the principle of term autonomy, the terms are listed

<sup>18</sup> The types of databases are discussed in detail in Tamás, Dóra-Mária: *A gazdasági szövegek fordításának terminológiai kérdéseiről*, Fordítástudományi értekezések I., Budapest, ELTE BTK, 2014.

separately, the entries are elaborate and include definitions, and the search results are presented in separate subject fields.

The criteria used in this study adapted the criteria used for the content analysis of print and electronic dictionaries<sup>19</sup> to the analysis of terminology databases, classifying the elements of mega-, macro-, micro- and mesostructure accordingly. In the case of Cercaterm, if we look at the megastructure, the search interface shows the description of the database as a subheading under the database title, and it says that the collection contains all the terminographic entries developed by Termcat. Clicking on the question mark icon in the search interface provides links to four short Catalan-language documents: a description of the database, a list of sources of terminological data, an explanation of abbreviations and language identifiers. The guide provides information on how to search and display the data and the number of entries; in 2018, the database contained 359,000 entries.

Turning to the macrostructure, the interface of the database can be explored in Catalan, Spanish, Occitan and English, with simple and advanced search modes. As regards the content of the database, Cercaterm provides access to the terms standardized by the Termcat Board of Supervisors through an alphabetical index in Catalan, Spanish, English, German, Italian and French, as well as to nomenclatures in Latin in the fields of botany and zoology. Using the simple search mode, users can search for a term or select a language from 45 languages, including Japanese, Sardinian and Guaraní. Of course, not all terms have equivalents in all of these languages, but the main term must be used in a search, and it is not possible to search by language or find out which terms are in the database in each language. As a result of the search, the main term always appears in Catalan in the list of results, with only some of the other language equivalents appearing in the entries. The advanced search can be used to specify key words and search the terms, the content of the definitions, the comments to the definitions, the language and the grammatical category of the items. There are terms and definitions belonging to 33 domains in a wide range of subject areas (e.g. health sciences, social sciences, sports, physics). If you search for a term using a simple search and select the language, you can click on the arrow in front of the term in the list of results to open the entry for that term. The entries contain only textual information and no images or multimedia.

In Cercaterm, the main term is Catalan. This gives the title of the datasheet, and then again the Catalan term is shown with its language identifier and grammatical category. The equivalents in other languages appear, also with

<sup>19</sup> See Fóris, Ágota – Rihmer, Zoltán: A szótárak minősítési kritériumairól, *Fordítástudomány*, Vol. 9, 2007/1, 109–113, and Gaál, Péter: Szempontrendszer online szótárak minősítéséhez, *Magyar Terminológia*, Vol. 5, 2012/2, 225–250.

their language identifier, followed by the subject heading, the Catalan definition and any comments on the definition. At the microstructure level, the structure of the entries will be illustrated by an example: a simple search for the Catalan term *península* (peninsula) in the search term field will return *península* as the only exact match in the list of results. Clicking on the arrow in front of the term opens the entry for the term, where the main term is in Catalan, followed by the relevant grammatical information (feminine noun), then the Spanish, French and English equivalents are listed, with a synonym under French. The subject field is indicated (*Geografia física, Geomorfologia*), as well as the Catalan definition of the term. Clicking on *Font* in the top right-hand corner of the entry opens the data sheet describing the source of the data in the entry.

In regard to the mesostructure, the system of references within the database features entries that are separated from each other and there are no internal links or links to external sources. An interesting feature of the entries is that they can be shared directly on Facebook and Twitter via the share icon, and a copiable link can be created.

### **Information on the use of Cercaterm**

The first thing to consider when looking at information on the use of the database is its usability and user-friendliness. In Cercaterm, the abbreviations system is logically structured, and clicking on the question mark icon in the search interface makes the list of abbreviations used and their explanations appear. The search interface is easy to use, user-friendly and free of superfluous, distracting elements, enabling searching to be quick and efficient. The regularity of data updates is also an important issue in terms of usability: in the case of Cercaterm, there is no precise date of the last update, but the database description says that it is updated monthly. The social utility of the database should also be considered: Cercaterm serves language policy purposes and is well-suited to achieving these purposes, as it contributes to the dissemination of terminological data as widely as possible.

#### ENTRIES ON THE TERM *Documentation* IN THE CERCATERM DATABASE

Fóris and Faludi, agreeing with Cabré's classification,<sup>20</sup> believe that the term *documentation* refers to technical writing, technical translation and technical documentation as well, while library documentation, also referred to as *docu-*

<sup>20</sup> Cabré, M. Teresa: *Terminology. Theory, methods and applications*, Amsterdam–Philadelphia, John Benjamins, 1999, 50.

*mentation*, is included in the field of document management.<sup>21</sup> The next part of the paper examines the content of the entries in the Cercaterm database that refer to the concept of documentation, the subject field of the term, and the Catalan definitions and equivalents in other languages that can be found in the database.

Writing the Catalan term for documentation (*documentació*) into the search window of the simple search function, when the language is set to Catalan and the subject field is set to all (*Totes*), yields the following results: 8 exact matches (*exactes*), 15 partial matches (*que conté*), and 180 approximate matches (*aproximat*s). Exact matches are displayed in a list, with the subject heading under the main Catalan term, and clicking on the arrow next to the term opens the entry for that term. On closer examination of the list, it is immediately apparent that there are three entries in which the term falls within the subject area of library science (*Biblioteconomia. Documentació*), so this list implies there is a case of homonymy. Clicking on the Source (*Font*) tab on the right-hand side of the first hit will open a window indicating that the entry needs to be revised and that the user should contact Termcat's customer support if they have any questions. Next comes the Catalan main term with the grammatical category (feminine noun) and the other language equivalents: Spanish, French, Italian and English, followed by the Catalan definition of the term. It defines documentation as the set of operations and techniques for the collection and processing of information, carried out in a continuous and systematic way, which enable the tasks related to the management of the documents generated during the documentation process to be carried out.<sup>22</sup> The definitions in the database are available only in Catalan. For the second entry, which also belongs to the subject area of library documentation, the source note also indicates that the entry needs revision. The main term is in Catalan, with equivalents in Spanish, French, Italian and English. The wording of the definition is different from the previous one, so it is clear that in this case it is a different concept. The definition says documentation is a set of collected or submitted documents.<sup>23</sup> The entry for the third concept belonging to the library documentation domain, designated with the same term, also includes the source note described above and equivalents in other languages, and the definition describes the term as a set of documents used to prove a fact, study or argument.<sup>24</sup>

<sup>21</sup> Fóris, Ágota – Faludi, Andrea: The Role of Documentation and Document Management in Translation and Terminology, in Fóris, Ágota – Bölcskei, Andrea (eds.): *Linguistic Research in the Fields of Content Development and Documentation*, Budapest–Paris, KRE–L'Harmattan, 2021, 139–155.

<sup>22</sup> “Conjunt d'operacions i tècniques de recollida i tractament d'informació, fetes de manera continuada i sistemàtica, que permeten de dur a terme les tasques de gestió de documents que descriu la cadena documental.”

<sup>23</sup> “conjunt de documents aplegats o adduïts”

<sup>24</sup> “aplec de documents que serveixen per a provar un fet, un estudi o una argumentació”

The next entry studied in this paper belongs to the subject field of document management (*Gestió documental*). The source of information is the Catalan language standard UNE-ISO 30300:2011 *Information and documentation – Management systems for records – Fundamentals and vocabulary*.<sup>25</sup> The ISO standard has been adopted as a Spanish standard by the Spanish Association for Standardization called AENOR,<sup>26</sup> and Termcat, the organization responsible for translating Spanish standards into Catalan, has ensured that the standard is also available in Catalan. An interesting feature of the entry is that it not only lists the Spanish, French and English equivalents of the term, but also includes a code number (3.1.4) which allows the definition to be easily located in the text of the standard. The definition of the term belonging to the subject field of document management is a set of documents describing operations, instructions, decisions, procedures or organizational rules related to a specific function, process or transaction.<sup>27</sup> In addition to the entries described above, terms belonging to the following subject areas were found: geography, cartography, criminal law, economics and corporate communications. In the subject field of cartography, the definition of the concept comes from the Termcat's electronic dictionary of geography. The entry on criminal law does not have a definition and also refers to an online Termcat thematic dictionary, which is accessible via a link. A common feature of all the entries is that the definitions all refer to documentation as a set of documents.

The partial hit (*que conté*) refers to multi-word expressions that contain *documentació* as an element of the compound. The analysis of partial results is not included in this study, and only some relevant aspects are described below. Among the results, no entry belongs to the subject area of document management, but several are found in the subject area of library science. The first example is the concept of *documentació factogràfica* (factographic documentation), which is defined as documentation relating to organizations, databases and experts working on terminology issues and terminology projects.<sup>28</sup> Other terms appearing among the results related to the subject area of library science include *centre de documentació de viatges* (documentation center for tourism), *documentació activa* (active documentation), *documentació semiactiva* (semi-active documentation), *documentació inactiva* (inactive documentation) and *servei de documentació* (documentation service). The entry for *servei de documentació* is the most extensive, and this is the only entry that contains an English equivalent. The source of the data is one of Termcat's printed thematic

<sup>25</sup> Informació i documentació: Sistemes de gestió per a documents: Fonaments i vocabulari

<sup>26</sup> <https://www.aenor.com/> (accessed June 30, 2020).

<sup>27</sup> “conjunt de documents que descriuen operacions, instruccions, decisions, procediments i regles organitzatives referents a una funció, un procés o una transacció determinats”

<sup>28</sup> “documentació que fa referència a organismes, a bancs de dades, a experts que treballen en qüestions terminològiques i als projectes terminològics en curs”

dictionaries, and the entry also includes synonyms in Catalan (*biblioteca especialitzada*, specialized library) and equivalents in other languages, including Spanish, French and English. Interestingly, the entry includes a synonym for each term in each language, with British and American versions of English terms. The definition of the concept of *servei de documentació* is as follows: an organization or service that manages documents stored on various media, distributes them and makes them available to users.<sup>29</sup> The definition is accompanied by a historical note: in the past, specialized libraries had traditional library holdings (*fondo bibliográfico*), but these institutions did not have access to databases, external catalogues or thematic research functions, so these tasks were carried out by documentation centers. Nowadays, both specialized libraries and documentation centers perform these tasks, and thus the choice between the two designations is a matter of individual institutional choice. With the additional explanation on the use of these synonyms, this entry is an excellent tool for understanding the meaning of the term. The third category of entries, the approximates (*aproximats*), is not described in detail within the scope of this study, as these concepts do not fall within the scope of either library science or document management, but belong to financial and economic fields.

Overall, three of the eight exact matches in the Cercaterm database for entries related to the concept of documentation belong to the subject area of library documentation. A comparison of the content of the terms reveals that they all refer to a set of documents, though the types of documents are different: the first term refers to procedures for operations, the second to documents certified in a more general sense, and the third to documents that provide evidence. Among the exact matches, one term falls within the scope of document management and its ISO standard definition is given in the entry. The partial results of the search do not include any terms from the field of document management, but several terms from the field of library science.

#### SUMMARY

Nowadays, new concepts and terms appear in large numbers. Delimitating, defining and designating concepts with appropriate terms are essential tasks for achieving clarity in professional communication. The creation and widespread availability of terminology databases contribute to standardizing terminology in various fields and improving the quality of professional documents and their translations. However, terminology databases not only have a direct practical benefit. High quality databases that are compiled according to ter-

<sup>29</sup> "organisme o servei que s'encarrega de gestionar documentació en qualsevol suport, de difondre-la i de posar-la a disposició dels usuaris"

minology principles also serve language policy purposes and fulfil terminology policy tasks. One such example is the Cercaterm database, edited by the Catalan terminology institute Termcat. This paper presented a detailed analysis of the database and described the entries for the term *documentation*. The Catalan government has paid close attention to the regulation of the official levels of language use, prioritizing the development of terminology because it had suffered decades of neglect over the last century. The database presented here is the result of a conscious effort by the professionals responsible for terminology; and the database as well as the associated workflow can serve as a good practice for other language communities.

## WORKS CITED

- BERTA, Tibor: A katalán az iberorómán nyelvek között, in Anderle Ádám (ed.): *Kutatói közlemények III. A magyar katalán kapcsolatok ezer éve*, Szeged, Hispania, 2001, 87–95.
- BERTA, Tibor: Nyelv és identitás: a katalán esete, *Szépirodalmi Figyelő*, 2018/1, 45–54.
- CABRÉ, M. Teresa: Terminology Today, in Somers, Harold (ed): *Terminology, LSP and Translation*, Amsterdam–Philadelphia, John Benjamins, 1996, 15–34.
- CABRÉ, M. Teresa: *Terminology. Theory, methods and applications*, Amsterdam–Philadelphia, John Benjamins, 1999.
- DÉRI, Balázs – FABULA, Kálmán: Fejezetek a katalán kultúra történetéből, *Szépirodalmi Figyelő*, 2018/1, 26–44.
- FÓRIS, Ágota – BÖLCSKEI, Andrea (eds.): Ajánlások a magyar terminológiastratégiához, in Fóris, Ágota – Bölcskei, Andrea (eds.): *Terminológiastratégiai kihívások magyar nyelvterületen*, Budapest, OFFI–L’Harmattan, 2019, 140–164, <https://www.offi.hu/offi-akademia/kiadvanyok/terminologiastrategiai-kihivasok-a-magyar-nyelvteru-leten>
- FÓRIS, Ágota – FALUDI, Andrea: The Role of Documentation and Document Management in Translation and Terminology, in Fóris, Ágota – Bölcskei, Andrea (eds.): *Linguistic Research in the Fields of Content Development and Documentation*, Budapest–Paris, KRE–L’Harmattan, 2021, 139–155.
- FÓRIS, Ágota – RIHMER, Zoltán: A szótárak minősítési kritériumairól, *Fordítástudomány*, Vol. 9, 2007/1, 109–113.
- GAÁL, Péter: Szempontrendszer online szótárak minősítéséhez, *Magyar Terminológia*, Vol. 5, 2012/2, 225–250.
- LÉNÁRT, András: Katalónia a 20. században, *RUBICON*, Vol. 28, 2018/11, 54–63.
- NEMES, Krisztina: Függetlenség – állam – nyelv: Katalónia különös esete, *Létünk*, 2015/2, 117–128.
- NÉMETH, Tibor: Regionális autonómia és nyelvpolitika Katalóniában, *Új Pedagógiai Szemle*, 2018/5–6, 116–122.

- PUJOL BERCHÉ, Mercè: Política lingüística: lengua cultura e identidad, el ejemplo de Cataluña, *Amnis, Revue d'études des sociétés et cultures contemporaines Europe-Amérique*, 2013/12, <https://journals.openedition.org/amnis/2061>
- SERMANN, Eszter: A dokumentáció terminushoz kapcsolódó fogalmak a Cercaterm katalán terminológiai adatbázisban, in Fóris, Ágota – Bölcskei, Andrea (eds.): *Tartalómfejlesztés és dokumentáció. Nyelvészeti kutatások*, Budapest, KRE–L'Harmattan, 2021, 247–262.
- SERMANN, Eszter – TAMÁS, Dóra: Hogyan definiálhatjuk a fordítói adatbázist? Egy olasz és egy spanyol fordítói terminológiai adatbázis vizsgálata, in Károly, Krisztina – Fóris, Ágota (eds.): *Nyelvek találkozása a fordításban. Doktori kutatások Klaudy Kinga tiszteletére*, Budapest, ELTE Eötvös, 2010, 101–116.
- SZENTE-VARGA, Mónika: Katalán nyelvpolitika, *Mediterrán világ*, Vol. 21, 2012, 53–63.
- TAMÁS, Dóra-Mária: *A gazdasági szövegek fordításának terminológiai kérdéseiről*, Fordítástudományi értekezések I., Budapest, ELTE BTK, 2014.
- TAMÁS, Dóra-Mária – SERMANN, Eszter: Elemzési szempontrendszer terminológiai adatbázisokhoz, *Fordítástudomány*, Vol. 21, 2019/2, 46–62.
- TAMÁS, Dóra-Mária – SERMANN, Eszter: Evaluation System for Online Terminology Databases, *Terminologija*, Vol. 26, 2020, 24–46.

## SOURCES

- AENOR, <https://www.aenor.com/>
- Cercaterm, <https://www.termcat.cat/ca/cercaterm>
- Generalitat de Catalunya, <http://web.gencat.cat/ca/inici>
- Institut d'Estudis Catalans, <https://www.iec.cat/>
- UNESCO: *Guidelines for Terminology Policies. Formulating and implementing terminology policy in language communities*, 2005, <https://unesdoc.unesco.org/ark:/48223/pf0000140765>
- UNESCO (Hungarian translation): *Terminológiapolitikai irányelvek. A terminológiapolitika kialakítása és megvalósítása különböző nyelvközösségekben* (original publication by: Infoterm, translated by Somogyi, Zoltán), in Fóris, Ágota – Bölcskei, Andrea (eds.): *Terminológiastratégiai kihívások a magyar nyelvterületen*, Budapest, L'Harmattan–OFFI, 2019, 167–218, <https://www.offi.hu/offi-akademia/kiadvanyok/terminologiastrategiai-kihivasok-a-magyar-nyelvterületen>
- Termcat. Centre de terminologia, <https://www.termcat.cat/ca>