

TERMINOLOGY IN MEDIA DISCOURSE: A CASE STUDY OF TERMS DENOTING PHOBIA TYPES IN ENGLISH, LITHUANIAN AND NORWEGIAN NEWS MEDIA SITES

SIGITA RACKEVIČIENĖ

Mykolas Romeris University, Vilnius, Lithuania
sigita.rackeviciene@mruni.eu

GIEDRĖ VALŪNAITĖ OLEŠKEVIČIENĖ

Mykolas Romeris University, Vilnius, Lithuania
gvalunaite@mruni.eu

KLAUDIJA CHEIKER

Western Union Processing Lithuania
klaudijac1@gmail.com

Abstract

The paper presents the trilingual (English – Lithuanian – Norwegian) analysis of the terms denoting phobia types in mass media discourse. The aim of the paper is threefold: to perform conceptual categorisation of the terms, establish the term formation patterns in the investigated languages, as well as to determine which phobia types were most often discussed in the selected news media sites (“The Guardian”, “DELFI” and “Dagbladet”) over a 10-year period. For the purposes of the research, a trilingual comparable corpus was compiled, from which 268 terms were manually extracted, matched and investigated. The findings of the research provide important information on conceptual, linguistic and social aspects of the phobia terms which may contribute to terminology research in the psychiatry domain.

Keywords: comparable corpora, conceptual categorisation, mass media discourse, phobia types, terminology, term formation patterns

1. Introduction

Terminology is the main tool of specialised communication, it enables to share knowledge in groups of specialists of various fields and transfer it to other people. Thus, terminology is inherent in specialised discourses: it is generated and

developed in specialised domains where it reflects the conceptual framework representing the knowledge accumulated by field experts.

However, terminology is also used outside specialised discourses; it enters the general language through various media channels, science popularization articles and discussions on the most relevant topics of the day. As Cabré (1999: 4) indicates, “The development of mass media communication allows the widespread dissemination of terminology, with the resulting interaction between the general and specialized lexicons. Specific terms become part of popular culture through their use in the mass media” (Cabré 1999: 4). Investigation of terms used in general discourse gives new insights into terminology: it reveals what specialised knowledge is most relevant at the given time period to the general public, how this knowledge is perceived, what terminological designations are preferred. Thus, investigation of terminology used outside specialised discourses gives significant additional information which can supplement terminology research in specialised domains.

The current article deals with the concept of phobia which was generated as a scientific concept of certain mental disorders in psychiatry domain. The terms denoting various phobia types have spread from the specialised psychiatry discourse to the mass media and through it to the general language. Some of them are widely known by the general public (*claustrophobia*, *aerophobia*, and *homophobia*); some still require more special knowledge (*arachnophobia*, *ophidiophobia*). The investigation of their use in mass media discourse allows collecting a lot of important terminological and social information which may contribute to the research in psychiatry domain.

The aim of the paper is threefold: to establish the conceptual categories of the terms denoting phobia types and linguistic means used for their formation, as well as to determine which phobia types were most often discussed in the selected news media sources. To achieve the aim the following objectives were pursued:

- 1) Reviewing the definition of concept ‘phobia’ in psychiatry domain;
- 2) Compilation of a trilingual comparable corpus of the selected articles from English, Lithuanian and Norwegian news media sites over a 10-year period and extraction of terms denoting phobia types;
- 3) Conceptual categorisation of the extracted terms;
- 4) Contrastive analysis of linguistic means used for term formation;
- 5) Thematic analysis of the articles aiming to determine which phobia types were most often discussed in the selected news media sites over a 10-year period.

The research is based on a comparable ad hoc corpus compiled for the purposes of the research which consists of 188 articles on phobia topics in English “The Guardian”, Lithuanian “DELFI” and Norwegian “Dagbladet” news media sites over a 10-year period (2009-2019).

2. Theoretical and methodological principles of the research

2.1. Terminology research and its dimensions

Modern terminology research is closely interrelated with cognitive science as it seeks to reveal how humans perceive the physical and abstract objects in the real world and their interactions, how this knowledge is conceptualised and structured in our minds and finally how constructed concepts are designated linguistically (cf. Cabre 1999, Faber 2009).

Construction of concepts and establishing their relations are the key processes in knowledge acquisition. They encompass several steps: we distinguish the essential characteristics of a class of similar referents and then by the process of abstraction construct an abstract mental unit representing the whole class. The process of conceptualisation enables us to categorise the objects of the real world into conceptual sets and to establish relations between different concepts inside the sets and among different sets. Our linguistic competence allows us to name the concepts and express ourselves. These categorisation and linguistic abilities are essential for our mental and social existence: “Because of the ordering of the concepts into sets, speakers are able to store a large amount of data and find them immediately when they require them to express themselves” (Cabre 1999: 39).

Thus, terminology research contributes to understanding of our thinking and cognition, knowledge acquisition and structuring patterns, as well as terminological designation of concepts and usage in communication.

Terminology work encompasses several dimensions which are determined by the multifaceted nature of terminological units:

- Cognitive dimension which relates linguistic forms (terms) with concepts representing classes of referents in the real world;
- Linguistic dimension which deals with the formation of terms in different languages;
- Communicative dimension which deals with use of terms in communication.

(cf. Sager 1990; Cabre 2003).

The cognitive dimension encompasses conceptualisation processes and organisation of concepts in the specialised domains. There are various criteria for organising concepts constituting the conceptual framework of the same domain. It depends on the nature of referents in the real world, as well as on the selected approach to categorise them. Terminologists distinguish two major types of relations among concepts:

- Logical: inclusive genus-species relations among concepts that share similar characteristics, but differ in the level of abstraction and are categorised as generic or specific towards one another in hierarchical classifications;

- Ontological: associative/thematic relations among concepts representing referents connected with each other in space, time, origin, interactional, cause/effect and other types of relations.

(cf. Nuoponnen 1994, 2014; Cabre 1999).

The establishment of conceptual relations enables to develop the whole conceptual framework of the specialised domain and/or elaborate its structural parts, as well as to develop systematic definitions of the meanings of terms.

The linguistic dimension encompasses the investigation of term formation patterns in different languages. It enables to reveal the use of internal/external lexical sources in concept designation process and formation methods of single-word and multi-word terms. The research of linguistic dimensions allows to disclose which term formation patterns prevail in different languages and gives new insights to term developers and translators (examples of multilingual research on linguistic dimension of terminology: Mockiene 2016; Smirnova and Rackevičienė 2018).

The communicative dimension focuses on pragmatic aspects of terminology: knowledge transfer in specialised communication and the role of terminology in it, as well as terminology translation and standardisation issues (Cabre 1999: 45–49).

Though the current research on phobia terms is based on its use in non-specialised mass media discourse and therefore cannot reveal the whole possessed knowledge of the domain, it encompasses all three dimensions of terminology investigation: conceptual categorisation of the collected terms, their linguistic analysis and insights on pragmatic aspects of the terminology.

2.2. Application of corpora in terminology research

Corpus linguistic has opened new possibilities for terminology. Digital corpora and corpus manager and analysis software allow working with big numbers of documents; extract comprehensive datasets necessary to examine cognitive, linguistic and communicative dimensions of terminology. Usually specialised ad hoc corpora are compiled for terminology research which represents the use of language in the selected specialised domain.

Bilingual and multilingual corpora are widely applied in contrastive research of terminology of several languages. They can be of two types: parallel or comparable.

The compilation principles of these two types of corpora differ essentially. Parallel corpora are composed of source language texts and their translations. Their compilation does not require any specific sampling frame, but needs thorough technical preparation, namely aligning the source and target texts, which is usually done on a sentence level (McEnery and Xiao 2007: 3).

Comparable corpora are composed of original texts in different languages which “deal with the same subject matter and were produced in similar situations of communication” (Delpech et al. 2012). Thus, the necessary prerequisite of

compilation of comparable corpora is the same sampling frame, i.e. “The components representing the languages involved must match with each other in terms of proportion, genre, domain and sampling period” (McEnery and Xiao 2007: 3). According to the degree of similarity, comparable corpora are further classified into “strongly comparable” and “weakly comparable corpora”: the former are “closely related texts reporting the same event or describing the same subject” (e.g. news items concerning the same specific event) and the latter are “weakly comparable texts which include texts in the same narrow subject domain and genre, but describing different events, as well as texts within the same broader domain and genre, but varying in sub domains and specific genres” (Skadiņa et al. 2010).

Parallel and comparable corpora have their own advantages and disadvantages. Parallel corpora allow revealing how the same content is expressed in different languages; they are especially important in translations studies as they allow observing what translation methods translators of different languages prefer to achieve semantic equivalence between the source and target languages. Machine-aided and machine translation systems benefit from having access to parallel data (McEnery and Xiao 2007: 4). However, parallel corpora “alone serve as a poor basis for cross-linguistic contrasts, because translations (i.e. L2 texts) cannot avoid the effect of translationese” (McEnery and Xiao 2007: 4).

Meanwhile, comparable corpora allow working with original languages which are much more natural. As Delpech et al. (2012) indicate that “Comparable corpora also have the advantage of containing more idiomatic expressions than parallel corpora do because the target texts do not bear the influence of the source language”. However, the matching process of the linguistic units is more complex and not all matches are possible to establish in comparable corpora.

Nevertheless, in the last decades comparable corpora (alone or in combination with parallel corpora) have been widely used in terminology. Various software solutions have been developed to apply them for bilingual terminology extraction (Vintar 2010; Gornostay et al. 2012; Aker, Paramite and Gaizauskas 2013). Their use widened the possibilities of terminology extraction and research as they allow collecting terminological data when parallel data do not exist or are not available.

For the present research, a comparable corpus of news media articles was compiled for terminology extraction and matching in the selected languages. According to the provided criteria, the compiled corpus might be regarded as partly strongly and partly weakly comparable: it consists of all phobia-related articles published in a 10-year period in the chosen media sites; some phobia types were discussed in all three media sites, some – only in two or one of them in the selected period. The choice of this type of corpora had several reasons: its sources were freely accessible, it allowed analysing phobia terminology in original languages and it enabled to reveal social aspects related to this mental disorder in three different countries.

3. Basic characterisation of the concept ‘phobia’ in scientific literature

The word *phobia* itself is derived from the Greek word *phobos* (φόβος), the original meaning of which is related to “fear, panic fear, terror, outward show of fear; object of fear or terror” (Harper 2020). However, it was not until the 19th century that it acquired its specialised meaning in the domain of psychiatry referring to mental disorder.

In the scientific literature, phobia is defined as “an irrational fear that produces a conscious avoidance of the feared object, activity or situation” (Buchanan and Coulson 2012: 3). Phobias are included into the classification systems of mental disorders where they are classified under the umbrella of anxiety disorders (Buchanan and Coulson 2012: 8). Twelve types of anxiety disorders are described in the Diagnostic and Statistical Manual of Mental Disorders published by American Psychiatric Association (2013). Several of them are specifically related to phobias and are referred to as phobic disorders. They comprise:

- social phobia (“persistent and intense fear of one or more social situations, due to a fear of showing anxiety symptoms or acting in an embarrassing way” (Buchanan and Coulson 2012)
- specific phobia (“excessive and persistent fear of a specific object or situation” (ibid))
- agoraphobia (agoraphobia is “anxiety about places or situations from which escape might be difficult, or in which help might not be available” (ibid).

Diagnostic criteria are developed to determine whether a person suffers from a phobic disorder and which type of a phobic disorder. However, these criteria are under constant debate (Buchanan and Coulson 2012). Various phobia types are also characterised in “The Encyclopaedia of Phobias, Fears, and Anxieties” (Doctor Kahn and Adamec 2008). The definitions of phobia types in the encyclopaedia were used for the conceptual categorisation of the extracted phobia terms.

4. Compilation of the trilingual comparable corpus and data extraction

4.1. Compilation of the corpus and its structure analysis

The trilingual comparable corpus, compiled for the purposes of the research, consists of the articles on phobia-related topics in three selected news media sites: English “The Guardian”, Lithuanian “DELFI”, Norwegian “Dagbladet”. The main criteria for extraction of articles were two:

- 1) the time of the article publication between 2009 and 2019;
- 2) the use of the term denoting the generic concept of phobia (*EN phobia*, *LT fobija*, *NO fobi*) and the complex terms denoting specific concepts referring to various phobia types and containing the key term *phobia/fobija/fobi* used

either as a part of a compound or a phrase (e.g. *EN orniphobia* = *LT ornitofobija* = *NO fobi for fugler*).

The compiled corpus consists of 188 articles extracted according to the above-described criteria. The total size of the corpus is 142 214 words. The sizes of the sub corpora are as follows: the English sub-corpus – 54839 words; the Lithuanian sub-corpus – 39415 words and the Norwegian sub-corpus – 47960 words. Figure 1 and Figure 2 present the distribution of the articles in the corpus.

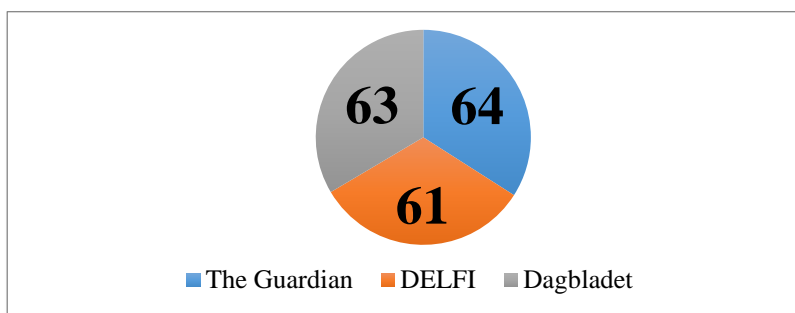


Figure 1. The number of articles extracted from each news media site

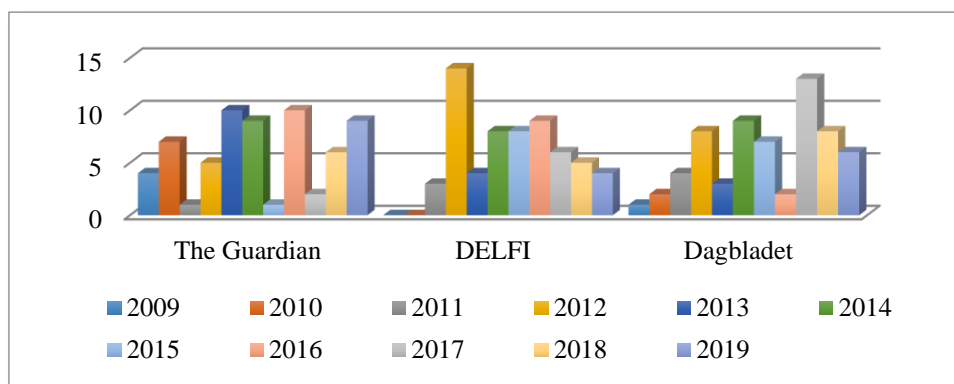


Figure 2. The distribution of the extracted articles over a 10-year period in each news media site

The data presented in Figure 1 show that each news media site shares almost an equal number of the extracted articles in the corpus. However, Figure 2 shows that the distribution of the extracted articles per year differs in each news media site. The highest number of the extracted articles was published in 2012 in “DELFI”, a total of 14 articles were published, followed by “Dagbladet” with 13 articles published in 2011, and “The Guardian” with 10 articles published both in 2013 and 2016. Also, in some years no articles on phobia topics were published at all. The corpus structure analysis reveals that phobia-related topics are almost equally relevant for the English, Lithuanian and Norwegian societies though the intensity of their discussions differ in different time periods.

4.2. Data extraction and systematising Heading B

The corpus contains various words with the lexical element *phob-/fob-/fob-* (e.g. *EN phobic*; *LT fobiškas*; *NO fobisk*), but as the main object of the research is the terms denoting phobia types, so only the terms referring to the specific concepts referring to various phobia types were extracted from the corpus. The data extraction was performed manually and the extraction from every sub-corpus consisted of the following steps:

- all terms denoting phobia types were manually detected, tagged and extracted from the articles;
- the extracted terms of each language were systematised on the conceptual basis (synonymic terms denoting the same concept were grouped together) and a list of terms was developed; the ratio of the concepts referring to phobia types and the terms denoting phobia types is presented in Figure 3:

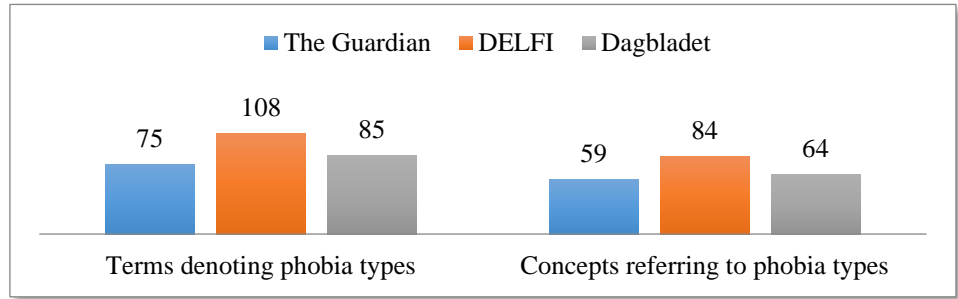


Figure 3. The ratio of the concepts referring to phobia types and the terms denoting them in each news media site

The figure reveals that the articles of “DELFI” discuss the largest number of phobia types; the second position was taken by “Dagbladet” and the third – by “The Guardian”. “DELFI” also included the largest number of synonymic terms.

The described data extraction and analysis was followed by manual matching of the terms in the three languages. Totally 157 concepts were distinguished and 14 of them had terminological designations in all three sub-corpora (English, Lithuanian and Norwegian), 20 concepts had designations in two of the sub-corpora and most 123 concepts had designations only in one of the corpora (see matching examples in Table 1)

Table 1. Examples of terminological designations of phobia types

| <i>EN</i> | <i>LT</i> | <i>NO</i> |
|--|--|--|
| <i>arachnophobia</i> <i>spider phobia</i> | <i>arachnofobija</i> <i>vorų fobija</i> | <i>edderkoppfobi</i> , <i>fobi mot</i> <i>edderkopper</i> , <i>fobi for edderkopper</i> |
| <i>trypophobia</i> <i>phobia of holes</i> | <i>tripofobija</i> | — |
| <i>ferry phobia</i> | — | — |

Thus, the comparable corpus only partly enabled to establish equivalents of the terms, but provided the sufficient and comprehensive material for terminological analysis of various aspects of the extracted data.

5. Conceptual categorisation of the extracted terms

In the first stage of the research, the extracted terms were categorised into conceptual categories based on logical and ontological relations among the concepts denoted by the terms. Firstly, 8 conceptual categories were distinguished according to the ontological relations among the concepts, i.e. according to the proximity of the referents, represented by the concepts, in the real world. Secondly, concepts of every category were further hierarchically subcategorised into more generic and more specific concepts into several hierarchical levels, the number of which depended on the characteristics of the concepts denoted by the terms in every category. The specific concepts of the same hierarchical level (in relationship of logical coordination) were further categorised according to their ontological relations.

Conceptual categorisation of the terms was based on their definitions in the specialised literature, the main sources of which was “The Encyclopaedia of Phobias, Fears, and Anxieties, Third Edition” (Doctor Kahn and Adamec 2008). The definitions enabled to distinguish 8 major conceptual categories relying on the nature of objects of different phobia types denoted by the extracted terms:

- Conceptual category of phobia types related to ecosystem;
- Conceptual category of phobia types related to man-made material things;
- Conceptual category of phobia types related to colours and symbols;
- Conceptual category of phobia types related to location and time;
- Conceptual category of phobia types related to physiology, health and death;
- Conceptual category of phobia types related to social situations;
- Conceptual category of phobia types related to groups of people and public figures;

- Conceptual category of phobia of everything (generalised fear).

The software “Free Mind” was used to visualise the structure of the established conceptual categories (Free Mind. Free software.). Due to limited space, only the major hierarchical levels of the categories are presented in the Figures below.

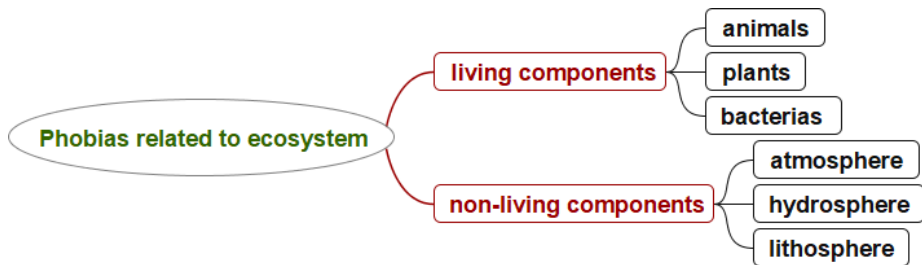


Figure 4. Conceptual category of phobias related to ecosystem

The conceptual category of phobias related to ecosystem (see Figure 4) contains the biggest number of terms detected in the investigated material. It was divided in two major subcategories “Phobias related to living components” and “Phobias related to non-living components”, and each of them was further divided into several groups and subgroups of more specific concepts.

The conceptual group “Animals” contains the biggest variety of subgroups and sub-subgroups: “Insects” (“Wasps”, “Butterflies”, “Bugs”, “Beetles”), “Arachnids” (“Spiders”, “Mites”), “Reptiles” (“Snakes” “Dinosaurs”), “Amphibians” (“Frogs”, “Toads”), “Birds”, “Mammals” (“Dogs”, “Cats”, “Mice”), “Worms” (“Tapeworms”). Several concepts “Animals” groups have terminological designations in all three sub-corpora (English, Lithuanian and Norwegian); other concepts have designations in one or two sub-corpora. The concepts denoted by the detected terms differ in their level of abstraction: some terms denoted the whole group of animals, some – certain species or subspecies of animals: e.g.

- the whole groups of animals: *EN animal phobia* = *LT zoofobija* / *fobija gyvūnams* = *NO fobi for ulike dyr, fobi for dyr; fobi mot dyr*;
- the species of animals *EN phobia of insects* = *LT vabzdžių fobija* = *fobi for insekter*;
- the subspecies of insects *NO fobi for veps* (‘phobia of wasps’);
- the subspecies of mammals: *EN musophobia* = *LT musofobija*.

The conceptual group “Plants” is much less numerous. In the corpus, the terms denoting two concepts representing two species of plants were detected, one of them has terminological designation in Lithuanian and another in Norwegian: *LT dendrofobija* (‘phobia of trees’), *NO anthofobi/ fobi mot blomster* / *blomsterfobi* (‘phobia of flowers’).

The conceptual group “Bacteria” contains one generic concept which has terminological designations in all three sub-corpora: *EN germ phobia* = *LT bakterijų fobija* = *fobi mot bakterier / bakterie-fobi*.

The subcategory “Phobias related to non-living components” contains a few concepts, none of them has terminological designations in all three sub-corpora. Most designations were detected in the Lithuanian sub-corpus, e.g.: *LT astrofobija/brontofobija/griaustinio fobija* (‘phobia of thunderstorms’), *homichlofobija* (‘phobia of fog’), *hidrofobija* (‘phobia of water’). One term was found in the Norwegian and one term in the English sub-corpus: *NO nephophobia* (‘phobia of clouds’), *EN phobia of mountains*.

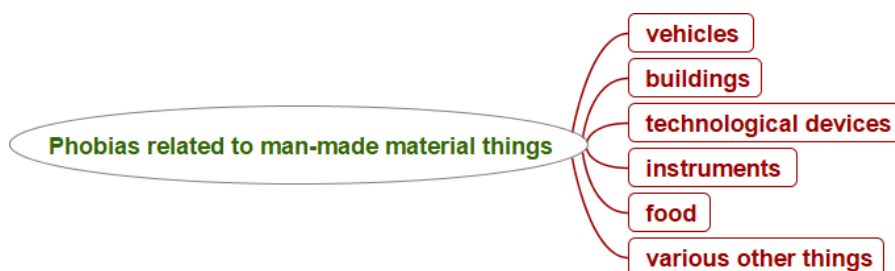


Figure 5. Phobias related to man-made material things

The category of phobias related to various man-made material things (see Figure 5) has several subcategories referring to various types of things manufactured or constructed by human beings. Most concepts of this category have designations only in one of the sub-corpora. The majority of them were detected in the English sub-corpus, e.g. *ferry phobia*, *phobia of tall buildings*, *mobile phobia*, *phobia of pens*, *phobia of bananas*, *white-wine phobia*, etc. In the Lithuanian and Norwegian sub-corpora the following terms were found: *LT automobilių fobija* (‘phobia of automobiles’), *NO fobi for elektriske gitarer* (‘phobia of electric guitars’), *NO consecotaleofobi* (‘phobia of chopsticks’), *NO lachanofobi* (‘phobia of vegetables’), etc. The concept ‘phobia of food’ has terminological designations in both English and Norwegian sub-corpora (*EN phobia towards food/phobia around trying new foods* = *NO matfobi*) and the concept ‘phobia of mirrors’ has terminological designations both in the Lithuanian and Norwegian sub-corpora (*LT eisoptrofobija* = *NO catoptrofobi*).

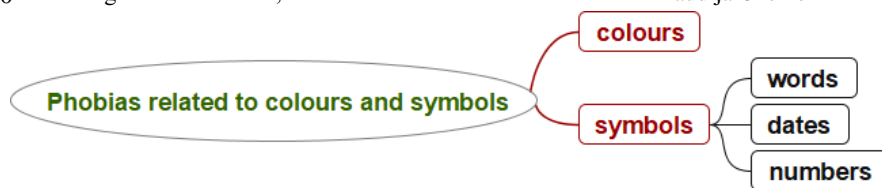


Figure 6. Phobias related to colours and symbols

The terms belonging to the conceptual category of phobias related to colours and symbols (see Figure 6) were found only in the Lithuanian and Norwegian sub-corpora, no English terms of this category were detected.

The Lithuanian and Norwegian terms designate the concepts referring to specific colours (yellow, red, black and white colours) or groups of colours (bright colours): *LT ksantofobija* ('phobia of yellow colour'), *LT fobija raudonai spalvai* ('phobia of red colour'), and *LT ryškių spalvų fobija* ('phobia of bright colours'), *NO melanofobi* ('phobia of black colour'). In both Lithuanian and Norwegian sub-corpora, the terms denoting the phobias of black colour were detected: *LT leukofobija*, *NO leukofobi*.

The terms denoting phobias related to symbols denote concepts referring to the number 13 (both in Lithuanian and Norwegian) and number 8 (only in Lithuanian): *LT paraskevidekatriafobija / frigatriskaidekafobija / penktadienio 13-os mėnesio dienos fobija*; *NO paraskevidekatriafobi*. In the Norwegian sub-corpus, the term denoting phobia of long words was detected: *NO hippopotomonstrosesquippedaliophobia*.

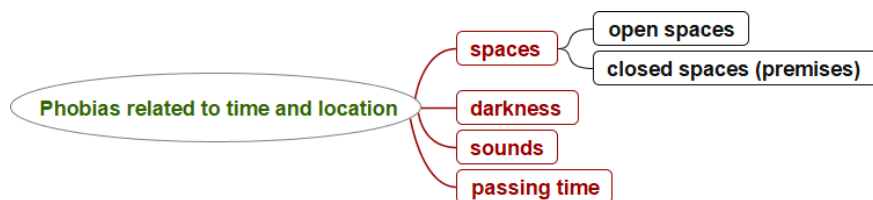


Figure 7. Phobias related to location and time

The conceptual category of phobias related to location and time (see Figure 7) encompasses mostly concepts representing open and closed spaces. Some of them are widely known and their terminological designations were found in all three sub-corpora:

- *EN claustrophobia* = *LT klaustrofobija / uždarų erdvių fobija* = *NO klaustrofobi / fobi for trange rom*;
- *EN phobia of heights / acrophobia* = *LT aukščio fobija / akrofobija* = *NO fobi for høyder*;
- *EN agoraphobia* = *LT atvirų vietų fobija / agorafobija* = *NO agorafobi* ('phobia of open or crowded places').

Other concepts had terminological designations in two sub-corpora or only in one of them: *EN phobia of sound* = *LT garsų fobija*; *LT niktofobija* / *tamsos fobija* = *NO fobi for mørke* ('phobia of darkness'). Only in the Lithuanian corpus, the term designating the phobia of passing time was detected: *LT chronofobija*.

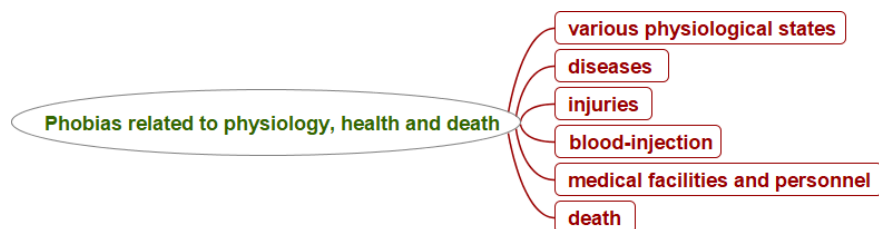


Figure 8. Phobias related to physiology, health and death

Another conceptual category of phobia concepts is related to human body (see Figure 8). This is a big group including various concepts referring to particular physiological states, diseases, injuries, death, medical facilities and personnel and other referents. Some concepts representing man-made material things and certain groups of people were also ascribed to this category as they are directly related to human health.

Most terms of this category were found in the Lithuanian sub-corpus, fewest – in the English sub-corpus. Only one concept (concept referring to phobia of needles) has terminological designations in all three sub-corpora: *EN needle phobia* = *LT adatų fobija* = *NO fobi for nåler*. Some terms denoting the same concepts were detected in two sub-corpora:

- *EN sickness phobia* / *phobia of being sick* = *LT nosofobija* / *nozofobija*;
- *EN phobia of hospitals* = *LT ligoninių fobija*.

However, most concepts had terminological designations only in one of the sub-corpora: *EN tokophobia* / *phobia of pregnancy*; *LT kancerophobia* ('phobia of developing cancer'), *LT senatvės fobija* ('phobia of getting old'), *NO blodfobi* / *fobi for blod* / *hemofobi* ('phobia of blood'), etc.



Figure 9. Phobias related to social situations

In the compiled corpus, a big group of terms denote concepts referring to various social situations related to social comfort and acceptability (see Figure 9). Some concepts represent fear of the situations in which people directly interact (*e.g. EN*

phobia of job interviews, phobia of kissing), some – represent fear of intellectual activities which can lead to unfortunate interaction (e.g. *LT desidofobija* ‘phobia of making decisions’) or fear of emotional states which can result from such interaction (e.g. *LT atazogorafobija* ‘phobia of being forgotten and ignored’).

The majority of the terms were found in the Lithuanian sub-corpus, the English terms took the second place, and the Norwegian sub-corpus contained fewest terms of this category. Some concepts of this category had designations in all three languages or two of them:

- *EN public speaking phobia / phobia about talking in public = LT glosfobija = NO lalofobi*;
- *EN commitment-phobia = LT įsipareigojimų fobija*;

Many terms did not have equivalents in other sub-corpora: *EN football phobia*; *LT katagelofobija* (‘phobia of being ridiculed’); *LT autofobija / isolofobija* (‘phobia of solitude’); *NO chorophobia* (‘phobia of dancing’); *NO epistemophobia* (‘phobia of acquiring knowledge’).

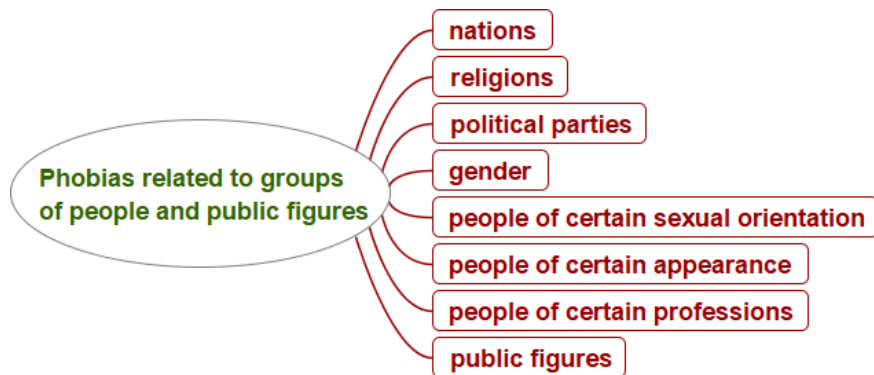


Figure 10. Phobias related to groups of people and public figures

The seventh conceptual category encompasses concepts referring to groups of people belonging to the same nation / religion / political party, etc. or certain specific individuals (see Figure 10). This category is special because many of its phobia concepts gradually develop additional property ‘prejudice/discrimination’, e.g. *EN Islamophobia* is defined as “irrational fear of, aversion to, or discrimination against Islam or people who practice Islam” (Merriam-Webster 2020).

Similar number of terms denoting the concepts of this category was found in all three sub-corpora. However, only two concepts had designations in all three sub-corpora:

- *EN homophobia = LT homofobija = NO homofobi*;
- *EN coulrophobia / clown phobia = LT kaurofobija = No klovne-fobi*.

Most terms did not have equivalents in other sub-corpora: *EN China-phobia*; *LT rusofobija / Kremliofobija / Rusijos fobija* (‘phobia of Russia’); *NO Frp-fobi*

(‘phobia of political party Fremskrittspartiet’), *NO Gretafofi* (‘phobia of Greta Thunberg’).

The last conceptual category is the category of phobia of everything. It contains two terms referring to the generic concept of generalised fear: *LT patofobija*, *NO panophobia*.

The conceptual categorisation of the extracted terms reveals the diversity of phobia objects discussed in the news media portals in three countries. These objects belong to various domains of the surrounding world: living/non-living components of ecosystem, man-made things, colours and symbols, locations and time, human body and health, social situations, particular individuals or groups of people or may even encompass everything. The concepts designated by the terms are of different degree of abstraction and represent bigger or smaller classes of referents in the real world.

The English terms denoted most concepts in the categories related to man-made material things and people (groups of people and public figures); the Lithuanian terms – concepts in the categories related to ecosystem, colours and symbols, location and time, human body (physiology, health and death) and social situations, while the Norwegian terms did not have anteriority in any of the categories.

6. Contrastive analysis of term formation patterns

In the second stage of the research, the formation patterns of the extracted terms were analysed and all terms were categorised according to two formation criteria:

- formal structure: compounds or phrases;
- source of the constituent denoting phobia object: internal or external lexical source.

The quantitative analyses were performed to reveal the major term formation patterns detected in the compiled corpus.

6.1. Formal structure of the terms

The terms denoting phobia types in all three languages are of two structural types: compounds and phrases.

In the English language three orthographical forms of compounds are distinguished: closed, hyphenated and open; all three forms were detected in the English sub-corpus: e.g. *musophobia*, *dino-phobia*, *bird phobia*. Meanwhile, in Lithuanian and Norwegian, compounds are of one orthographic form – closed compounds, they were detected both in the Lithuanian and Norwegian sub-corpora, e.g. *LT zoofobija* ‘phobia of animals’; *NO blomsterfofi* ‘phobia of flowers’.

The phrases denoting phobia types in the corpus are nominal phrases in which the word *phobia/fobija/fobi* functions as the head of the phrase in all three languages. The modification patterns and syntactic structure of the phrases differ in the languages. In English and Norwegian, the heads of the phrases are post-modified by prepositional phrases, e.g. *EN phobia of school*, *phobia about being tall*; *NO fobi for slanger*, *fobi mot slanger* ‘phobia of snakes’. In Lithuanian, on the other hand, the heads of the phrases are pre-modified by nominal phrases in the Genitive or Dative cases, e.g. *automobilių fobija* ‘phobia of automobiles’, *fobija raudonai spalvai* ‘phobia of the red colour’.

Many concepts, distinguished in the research, are denoted by terms of different structural types (or, in English, different orthographic forms), e.g.

- *EN arachnophobia* / *spider phobia* = *LT arachnofobija* / *vorų fobija* = *NO edderkoppfobi* / *fobi mot edderkopper* / *fobi for edderkopper*.
- *EN acrophobia* / *phobia of heights* = *LT akrofobija* / *aukščio fobija* = *NO fobi for høyder*.

The quantitative analysis of the formal structure of the terms reveals that the ratio of compounds and phrases is similar in all three languages (see Figure 11): the compounds clearly prevail over the phrases in the investigated languages. In the Lithuanian and Norwegian datasets the ratio is identical whereas in the English dataset, the number of the compounds is lower by 5 % and number of phrases is higher by 5 % in comparison to the Lithuanian and Norwegian data.

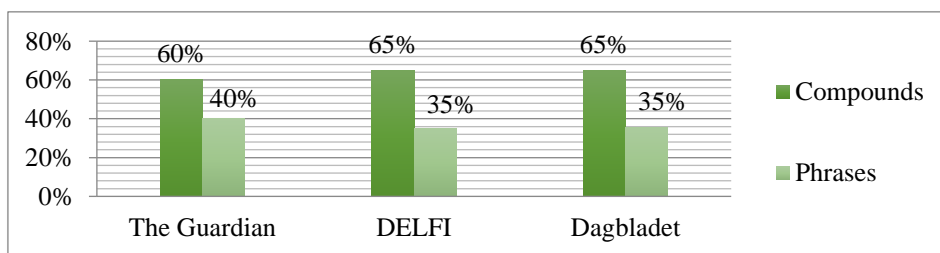


Figure 11. The percentage of the terms according to their formal structure

6.2. Lexical sources of the object-denoting constituent of the terms

The terms denoting phobia types consist of constituents of different origin. The constituent *phobia/fobija/fobi* is of Greek origin in all three languages. Meanwhile the object-denoting constituent may be formed using internal or external lexical sources. In the present research, internal sources encompass both words of native origin and old borrowings fully assimilated into the languages and common in general use of the languages, e.g. *EN animal*, *mountain*, *school*.

In the compounds, the object-denoting constituent is formed using either external or internal lexical sources. The most usual external source is the Greek language, followed by Latin; other external sources are used only in few cases.

The internal sources for object-denoting word are used only in the English and Norwegian languages:

- Examples of use of external sources:

Greek origin: *EN agoraphobia* = *LT agorafobija* = *NO agorafobi* ‘phobia of open or crowded places’; *LT triskaidekafobija* ‘phobia of the number 13’; *NO epistemophobia* ‘phobia of acquiring knowledge’.

Latin origin: *EN claustrophobia* = *LT klaustrofobija* = *NO klaustrofobi*.

Arabic origin: *EN Islamophobia* = *NO Islamofobi*.

- Examples of use of internal sources:

EN spider phobia = *NO edderkoppfobi*; *EN fatphobia*; *NO blodfobi* ‘phobia of blood’.

In the phrases, the object-denoting constituent is taken mostly from internal lexical sources. Use of external sources is detected only in few cases:

- Examples of use of internal sources:

EN phobia of snakes = *LT gyvačių fobija* = *NO fobi for slanger*; *EN phobia of tall buildings*; *LT tamsos fobija* ‘phobia of darkness’; *NO matfobi* ‘phobia of food’.

- Example of use of external sources:

Latin origin: *NO fobi for transpersoner* ‘phobia of transgender people’;

Armenian origin: *EN phobia of the Kardashians*.

The quantitative analysis of the lexical sources of the object-denoting constituent of the terms denoting phobia types reveals that the use of internal and external lexical sources for formation of the object-denoting constituent is rather different in the languages (see Figure 12). In the English dataset, the use of internal sources obviously prevails while in the Lithuanian dataset, on the contrary, the use of external sources dominates. In the Norwegian dataset, the numbers of terms that include the object-denoting constituent formed using internal sources and external sources are almost equal with slight predominance of external sources.

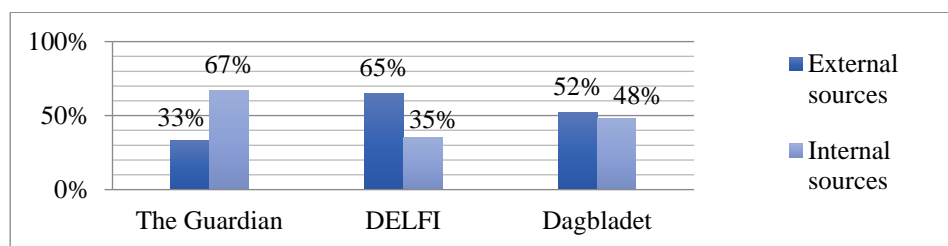


Figure 12. The number of the terms according to the sources of object-denoting constituent

7. Contrastive thematic analysis of the news media articles

In the last stage of research, the thematic analysis of the news media articles was performed aiming to determine which phobia types were most often discussed in the articles in the period 2009-2019. The results of the quantitative analyses of the articles are presented in the Figures 13, 14, 15.

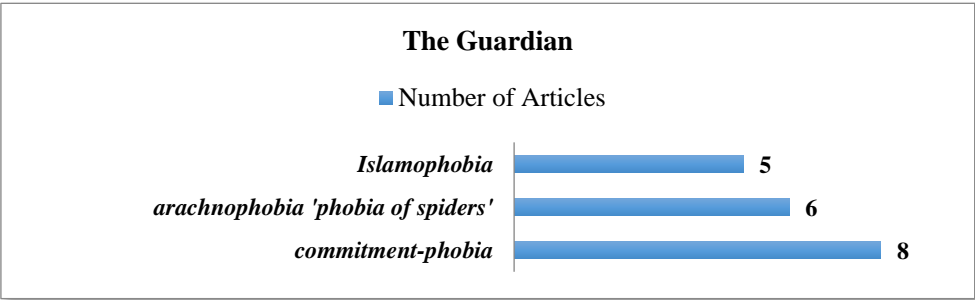


Figure 13. TOP 3 phobia types according to the number of articles discussing them in “The Guardian”

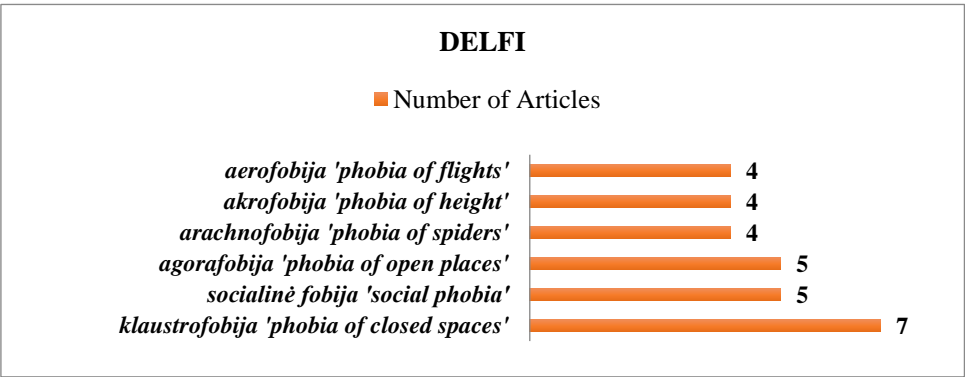


Figure 14. TOP 3 phobia types according to the number of articles discussing them in “DELFI”

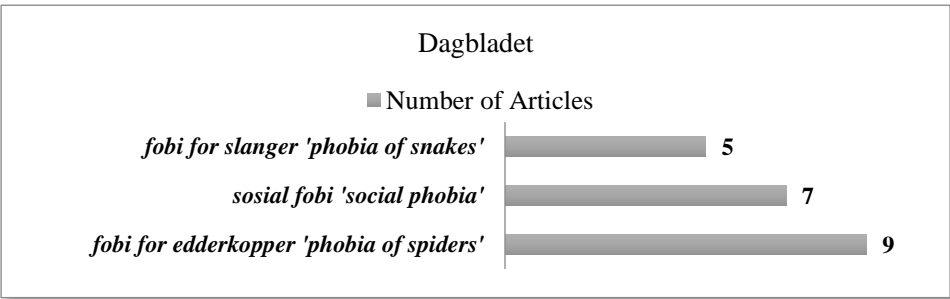


Figure 15. TOP 3 phobia types according to the number of articles discussing them in “Dagbladet”

The quantitative analysis reveals that in the Guardian the most discussed phobia was the phobia ascribed in the research to the conceptual category of phobias related to social situations, denoted by the term *commitment-phobia*, 8 articles (13 % of all phobia-related articles) discussed it in the time period 2009-2019. In the Lithuanian news media site, the most discussed phobia belongs to the conceptual category of phobias related to location and time and is denoted by the term *klaustrofobija* ‘claustrophobia’, 7 articles (11 % of all phobia-related articles) discussed it in 2009-2019. Meanwhile, in the Norwegian news media site, the most discussed phobia conceptually belongs to the phobias related to ecosystem, it is denoted by the term *fobi for edderkopper* ‘phobia of spiders’, 9 articles (14 % of all phobia-related articles) discussed it in the investigated period.

The analysis also reveals that phobia of spiders is among the most discussed phobia types in all three news media sites (9 %, 7% and 14 % of all phobia-related articles in the English, Lithuanian and Norwegian media sites respectively), while social phobia is among the most discussed phobia types in the Lithuanian and Norwegian media articles (8 % of all phobia-related articles in the Lithuanian media site, and 11 % in the Norwegian media site).

8. Conclusions

The findings of the research allow drawing several important conclusions.

The analysis of the structure of the compiled corpus shows that the number of the phobia-related articles extracted from the media sites is quite similar (64, 61 and 63 articles from the English, Lithuanian and Norwegian media sites respectively), though the distribution of the articles over the selected period (2009-2019) differs in each news media site. This shows that phobia-related topics are equally relevant for the societies in all three countries.

The number of the extracted terms from different media sites differs more considerably, it reveals that the Lithuanian media site discussed the biggest variety of phobia types (84) in the selected time period, the second position was taken by the Norwegian media site (64), and the last – by the English media site (59). In the whole corpus, 157 concepts representing different phobia types were distinguished. 9 % of them had terminological designations in all three sub-corpora (English, Lithuanian and Norwegian), 13 % of the concepts had designations in two of the sub-corpora and even 78 % of the concepts had designations only in one of the corpora. The term-matching results disclose that comparable corpora only partly enable to find equivalent terms in different languages.

The conceptual categorisation of the extracted terms reveals the diversity of phobia objects discussed in the news media sites in three countries. These objects

belong to various domains of the surrounding world: living/non-living components of ecosystem, man-made things, colours and symbols, locations and time, human body and health, social situations, particular individuals or groups of people or may even encompass everything. The English terms denoted most concepts in the categories related to man-made material things and people; the Lithuanian terms – concepts in the categories related to ecosystem, colours and symbols, location and time, human body and social situations, while the Norwegian terms did not have anteriority in any of the categories.

The analysis of the linguistic means used for term formation reveals the following points:

- the terms denoting phobia types are formed either as compounds or phrases; the compounds prevail over phrases in all three languages;
- the constituent denoting the object of the phobia, can be formed using either internal or external sources; in the compounds, the object-denoting constituent is formed using either external (predominantly Greek) or internal sources while in the phrases the use of internal sources significantly dominates;
- in the English dataset, the use of internal sources prevails while in the Lithuanian dataset, on the contrary, the use of external sources dominates. In the Norwegian dataset, the numbers of terms that include the object-denoting constituent formed using internal sources and external sources are almost equal.

The thematic analysis of the articles devoted to different phobia types shows that the most often discussed phobia types in the investigated news media sites are as:

- in “The Guardian”: *commitment-phobia* ‘phobia of committing to someone’;
- in “DELFI”: *klaustrofobija* ‘phobia of closed spaces’;
- in “Dagbladet”: *fobi for edderkopper* ‘phobia of spiders’.

Phobia of spiders is among the TOP 3 phobia types in all three news media sites, while *social phobia* ‘phobia of social situations’ is among TOP 3 phobia types in “DELFI” and “Dagbladet”.

The findings of the research also allow making some important methodological conclusions. The research shows that the investigation of terminology used outside specialised discourse enables to reveal significant additional information about terminology which may contribute to its research in the specialised domain: what specialised knowledge is most relevant to the general public at a particular time, what knowledge units are discussed in different societies, what social, cultural, political and other factors have impact on perception of these knowledge units and their development, etc. The research also shows that the use of comparable corpora enable to perform multidimensional contrastive analysis and in addition to conceptual and linguistic information reveal various social aspects of terminology use in different societies which cannot be revealed by parallel corpora.

References

- Aker, Ahmed; Paramite, Monica and Robert Gaizauskas. Extracting bilingual terminologies from comparable corpora. 2013. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, August 4-9, Sofia, Bulgaria*. 402-411.
- American Psychiatric Association. 2013. *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)*. Arlington, VA: American Psychiatric Association. <https://doi.org/10.1176/appi.books.9780890425596>
- Buchanan, Heather and Neil Coulson. 2012. *Phobias* (Series *Palgrave Insights in Psychology*). Palgrave Macmillan. <https://doi.org/10.1007/978-0-230-36370-0>
- Cabré Castellví, Maria Teresa. 1999. *Terminology: Theory, Methods and Applications*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Cabré Castellví, Maria Teresa. 2003. Theories of terminology: their description, prescription and explanation. *Terminology* 9 (2). 163-199. <https://doi.org/10.1075/term.9.2.03cab>
- Delpech, Estelle; Daille, Béatrice; Morin, Emmanuel and Claire Lemaire. 2012. Extraction of Domain-Specific Bilingual Lexicon from Comparable Corpora: Compositional Translation and Ranking. *COLING 2012*. Mumbai, India. Available from <https://arxiv.org/abs/1210.5751> [Accessed: 6th July 2020].
- Doctor, Ronald M.; Kahn, Ada P. and Christine A. Adamec. 2008. *The Encyclopaedia of Phobias, Fears, and Anxieties*. Third Edition New York: Fact on File Inc.
- Faber Benítez, Pamela. 2009. The Cognitive Shift in Terminology and Specialized Translation. *MonTI. Monografías de Traducción e Interpretación* 1. Universitat de València. 107-134. <https://doi.org/10.6035/MonTI.2009.1.5>
- Free Mind. Free software. Available from: http://freemind.sourceforge.net/wiki/index.php/Main_Page
- Gornostay, Tatiana; Ramm, Anita; Heid, Ulrich; Morin, Emmanuel; Harastani, Rima and Emmanuel Planas. 2012. Terminology Extraction from Comparable Corpora for Latvian. *HLT 2012: 5th International Conference of Human Language Technologies, October 2012, Estonia*. 66-73.
- Harper, Douglas. 2020. Entry for 'phobia.' *Online Etymology Dictionary*. Available from: <https://www.etymonline.com/search?q=phobia> [Accessed: 6th July 2020].
- McEnery, Anthony and Zhonghua Xiao. 2007. Parallel and comparable corpora: What are they up to? In Gunilla Anderman and Margaret Rogers (eds.) *Incorporating Corpora: Translation and the Linguist* (Series *Translating Europe*), chapter 3, 18-31. Clevedon: Multilingual Matters. <https://doi.org/10.21832/9781853599873-005>
- Merriam-Webster. 2020. Entry for 'Islamophobia'. *Merriam-Webster*. Available from: <https://www.merriam-webster.com/dictionary/Islamophobia> [Accessed: 6th July 2020].
- Mockienė, Liudmila. 2016: *Formation of terminology of constitutional law in English, Lithuanian and Russian*. Doctoral Thesis, Vilnius: Mykolas Romeris University.
- Nuopponen, Anita. 1994. *Begreppssystem för terminologisk analys* [Concept systems for terminological analysis]. PhD dissertation. Acta Wasaensia. University of Vaasa, Vaasa.
- Nuopponen, Anita. 2014. Tangled Web of Concept Relations. Concept relations for ISO 1087-1 and ISO 704. *Terminology and Knowledge Engineering, TKE 2014*, Berlin. Available from: <https://hal.archives-ouvertes.fr/hal-01005882/document> [Accessed: 6th July 2020].
- Sager, Juan Carlos. 1990. *A Practical Course in Terminology Processing*. Amsterdam/Philadelphia: John Benjamins Publishing Company. <https://doi.org/10.1075/z.44>
- Skadiņa, Inguna; Aker, Ahmet; Giouli, Voula; Tufis, Dan; Gaizauskas, Rober; Mieriņa, Madara and Nikos Mastropavlos. 2010. A Collection of Comparable Corpora for Under-resourced Languages. In Skadiņa, Inguna and Andrejs Vasiljevs (eds.) *Human Language Technologies – The Baltic Perspective: Proceedings of the 4th International Conference* (Book Series *Frontiers in Artificial Intelligence and Applications*, vol. 219). 161 – 168.

- 380 Sigita Rackevičienė, Giedrė Valūnaitė Oleškevičienė and Klaudija Cheiker
Smirnova, Oksana and Sigita Rackevičienė. 2018. Corpus-Driven Analysis of Multi-Word Terms Including the Word 'Risk' in English, French and Lithuanian. *Terminologija* 25. 86–106.
Vintar, Špela. 2010. Bilingual term recognition revisited: The bag-f-equivalents term alignment approach and its evaluation. *Terminology* 16 (2). 141-158.
<https://doi.org/10.1075/term.16.2.01vin>

Data sources (as of 2019)

The Guardian. Available from: <https://www.theguardian.com/international>
DELFI. Available from: <https://www.delfi.lt/>
Dagbladet. Available from: <https://www.dagbladet.no/>