

Citation: Łukasik, M. (2017). Contrastive terminography. *Cognitive Studies / Études cognitives*, 2017(17). <https://doi.org/10.11649/cs.1378>

MAREK ŁUKASIK

Pomeranian University in Słupsk, Słupsk, Poland
marek.lukasik@apsl.edu.pl

CONTRASTIVE TERMINOGRAPHY

Abstract

Contrastive methods have long been employed in lexicography, in particular in bi- and multilingual dictionary projects. The main rationale for this is the necessity to comprehensively study, i.e. compare and contrast, two or more linguistic systems that are to be presented in one way or another in respective dictionaries. Similarly, the contrastive approach is of paramount importance in terminographic undertakings, on account of the need to draw a distinction between terminological (conceptual) systems existing in various languages and across cultures. It must be emphasised, however, that the contrastive element is not only a part of terminographic practice, but also of the theory of terminography. This article aims to present the role of contrastive research across various spheres of specialised (=LSP) lexicography.

Keywords: comparative lexicography; contrastive specialised dictionary; conceptual analysis; terminological studies; terminographic analysis; parallel corpus; comparable corpus

1 Introduction

Contrastive studies have traditionally been associated with the compilation of bilingual and multilingual dictionaries, mainly on account of the fact that such works present linguistic units belonging to at least two, often divergent, language systems. Accordingly, the major object of analysis in such works has been equivalence (Hartmann, 1991, p. 2854).¹ However, the relation between lexicography and contrastive studies definitely goes beyond this simple relationship. Intra- and interlingual dictionaries are (or can be) the means of presenting a whole range of other results of contrastive studies that might be of interest to a wide range of specialists. For example, special-purpose dictionaries of language universals, borrowings, or dialectal variations could prove highly useful resources for linguists, while translators could greatly benefit from dictionaries of false-friends, idioms and — yet to be fully realised — cross-cultural dictionaries. Interestingly enough, monolingual and bi-/multilingual dictionaries — if compiled in accordance with best lexicographic practice — can themselves be a reliable source for further contrastive analyses.

¹It needs to be emphasised, however, that a bilingual dictionary is not automatically a (type of) comparative/contrastive dictionary; only in specific circumstances, and by following appropriate contrastive analyses, can a bilingual dictionary meet the requirements of a comparative/contrastive reference work.

Definitely, one of the most obvious lexicographic products which present the results of contrastive studies are comparative dictionaries. As the *Dictionary of Lexicography* holds:

[a] comparative dictionary [is] a type of dictionary which is compiled, sometimes on the basis of informal glossaries collected during linguistic fieldwork, for the purpose of contrastively evaluating the shared vocabulary of several languages. (DoL, 2002, p. 24)

On the one hand, the definition presented is broad enough to allow for a whole array of types of dictionaries that could be subsumed under the umbrella term considered. Undoubtedly, each of the dictionary types will differ as regards to the aim of the vocabulary evaluation, and in terms of the *tertium comparationis* used. On the other hand, the definition is restrictive, as it limits the object of analysis to vocabulary only. Given the great number of research areas in contemporary linguistics, such a limitation seems to be unsubstantiated.

The perspective presented above concerns only one side of lexicography, namely the practical enterprise of dictionary-making. Lexicography, however, has also developed its theoretical underpinnings, consisting of systematic dictionary research, dictionary criticism, research into dictionary use and user needs, studies of the dictionary market, and the drafting of the history of lexicography (Hartmann, 2001, pp. 6–8, 27–30; Piotrowski, 1994, p. 11). It can easily be envisaged that at least in some, if not most, of the areas of metalexicographic research, contrastive studies could, and/or should, be universally applied. Possibly on account of the fact that theoretical studies in lexicography have recently gained traction, and in light of the need to unify lexicographic methods, some scholars have perceived the need for a new area of lexicographic endeavour termed ‘comparative’. The definition of comparative lexicography, as presented in *DoL*, reads:

1. A branch of general lexicography, which contrasts the dictionary traditions of various cultures, languages and countries with a view to distilling from them common principles by considering the external factors that have led to divergent practices. Examples include issues such as how different scripts influence the format of reference works, which different genres predominate, and what constitutes good practice in DICTIONARY MAKING and dictionary use.
2. A complex of activities concerned with the design, compilation, use and evaluation of comparative dictionaries. (DoL, 2002, p. 24, emphasis as in the original)

It clearly transpires from the definition (1) above that contrastive analysis is no longer treated solely as the means of establishing and differentiating (disambiguating) meanings in bilingual dictionaries (i.e. by way of equivalence analysis, as outlined above) or representing discrepancies between languages in such works (a broadened approach to the comparative dictionary), but rather can be seen as an umbrella method used for a plethora of other studies, including (meta)lexicographic research. This would seem to be confirmed by the second part of the definition (2), which in fact prescribes a whole range of tasks linked to dictionary-making and metalexicographic analyses, albeit limited to comparative dictionaries only. It is worth emphasising at this point that contrastive analyses involve two steps: description (of the source data) and analysis (James, 1980, p. 63). Only then can the studies produce reliable results of an applicable nature.

2 The contrastive approach to metaterminography

Considered by some scholars to be a branch of lexicography, terminography is the theory and practice of specialised dictionary-making. As with lexicography, contrastive studies may be applied to both its theoretical component and to the preparation of specialised (terminological) dictionaries (cf. the definitions above). The following sections will discuss the specific areas of terminography in which the contrastive component could, and/or should, definitely be applied.

2.1 General terminographic analysis

Since most new dictionaries are — to a lesser or greater extent — based on existing works, i.e. on their structural composition, compilation methodology, scope and, in some instances, even content — systematic dictionary studies (the description phase of terminographic endeavours) are a necessary step towards ensuring high quality lexicographic work. The aim of such research is to distil successful lexicographic solutions, and minimise the spread of faulty constructions and factual errors. The qualities most sought for in lexicography are high accuracy of linguistic content and a better response to users' needs. As Saloni emphasises, one of the means to achieve these qualities is multi-level, comprehensive analysis of dictionaries (Saloni, 1988, pp. 8–9). It is high time that theoretical considerations started to be regarded in lexicography and terminography as the necessary initial step which precedes the dictionary-making process.

However, it is not until recently that such systematic studies have been undertaken in respect to terminological dictionaries published in Poland (Łukasik, 2007). Terminographic analysis, as this type of study is known, is geared towards research into existing specialised dictionaries (and other related reference works produced by terminographers), and can take the form of general or detailed terminographic analysis. The overarching goal of general terminographic analysis is the assessment of the state of terminographic practice in a given area (such as, for example a country) or of a particular society (such as a nation) over a particular time. Detailed terminographic analysis investigates dictionaries according to a complete or pre-selected set of parameters at all structural levels (mega-, macro-, medio- and microstructure) as well as content-wise. In fact, full terminographic analysis is only possible following detailed studies of existing terminographic works (Łukasik, 2012, p. 102). In any case, setting beside two or more lexicographic works (=carriers of linguistic data) entails some form of contrastive analysis.

By applying the definition of comparative lexicography to terminography, it becomes clear that general terminographic analysis outlines a methodology that can also be employed when undertaking the comparison of entire national terminographic systems (i.e. tradition, methodologies, practices, etc.). Some form of contrastive analysis seems necessary to compare (and contrast) such systems across various cultures or languages. It needs to be emphasised here that the studies may involve two distinct terminological systems within one language, as in the case of legal terminology (for example, the American vs. the British legal system). An example of an attempt to apply comparative terminographic analysis is the study of bi- and multilingual specialised (= LSP) dictionaries of Polish and Ukrainian terminology, published in Poland and Ukraine. This study has provided an insight into, and a comparison of, the terminographic practices of the two countries, as well as the identification of areas that require joint terminographic efforts (Łukasik, 2016a, 2016b).

2.2 Detailed terminographic analysis

Since detailed terminographic analysis involves the study of individual dictionaries across a variety of parameters, which may amount to dictionary criticism if a work is compared to a model or another published dictionary, contrastive methods — especially in relation to the dictionary content — come to the forefront, regardless of whether the researcher deals with mono- or bi-/multilingual work. Since terminological dictionaries are carriers of terms, i.e. language representations of concepts, such comparisons equally concern the linguistic means and the conceptual systems reflected in particular dictionaries. As regards different terminological systems, the analysis involves a reconstruction and comparison of two or more conceptual systems encoded in the terminology included in the dictionary. The conceptual network behind the primary terminological system, i.e. the system in which the professional knowledge is represented to the highest degree of precision (Lukszyn & Zmarzer, 2006, p. 60 ff.), and more specifically the set of characteristics connected with each concept — becomes the *tertium comparationis* in such analyses. The starting point here is the listing of concept features in each of the conceptual systems analysed, their comparison to

the respective set of features in the primary terminological system, as well as an evaluation of the degree to which the systems overlap, both quantitatively (if feasible) and qualitatively. The contrastive/comparative linguistic studies of terms comprise the language rules of each of the languages studied, including the rules of intra- and interlingual borrowings, as well as terminology internationalisation principles. The invariable element that ensures a high-quality comparison is the conceptual layer of the concept-term pair.

A study by de Groot and van Laer (2008) is a good example of a detailed dictionary study, in which the terminographic comparison method was used. The authors undertook a multidirectional assessment of approximately 200 legal dictionaries published in the EU. Interestingly enough, the authors emphasise that it was not difficult to make a list of very poor, even dangerous (sic!) bilingual dictionaries in the field, with only 12 out of 200 dictionaries being of high quality (de Groot & van Laer, 2006, p. 80). Not surprisingly, the main issue in the opinion of the authors cited is the lack of comparative meaning explanation in the works analysed, which is a matter of primary importance in presenting culture-dependent terminology. There are numerous legal systems not covered by the existing dictionaries and commercial publishers keep continue to sell bilingual dictionaries of inferior quality because there are no other translation tools for the language pair concerned. In fact, most dictionaries are wordlists, i.e. bi- or multilingual lists of terms offering unsubstantiated translations, with no explanation as to the different meanings offered (de Groot & van Laer, 2006, p. 75). Pieńkos (2003) neatly explains the reasons for this situation:

Globally, studies in the field of comparative legal lexicography have not yet achieved a level which would enable the compilation of comprehensive dictionaries or encyclopaedias — ones that would be complete and good enough to be considered useful guidelines to legal terminology.

(Pieńkos, 2003, p. 311^a)

^aTranslation mine, M.Ł.

This short summary, viewed from the large-scale perspective of terminographic theory and practice, constitutes a strong desideratum as to the directions of modern terminographic development.

2.3 Other metaterminographic studies

Other areas of research in metaterminography that can benefit from the application of contrastive methods include dictionary user/use research, or studies into the history of terminography. The former aims to identify the profile of potential dictionary users and to establish dictionary functions tailored to users' needs. Contrastive studies seem best to fit user research projects in which the user group is linguistically heterogeneous and in which some common linguistic traits need to be distilled, as well as in cases when idiolectal differences need to be taken into account (as in author dictionaries).

Researching the history of terminography is inextricably connected to synchronic or diachronic, intra- or interlingual contrastive studies, chiefly because these studies necessitate the analysis of various types of terminographic works. Besides the basic and general objective of summarising the achievements (tradition) of humanity (society) in terms of the presentation of (a portion of) terminological material, the ultimate goal of historical studies can be practical: the compilation of ever better terminographic works in the future, based on research into successful and failed terminographic undertakings from the past.

3 The practice of contrastive terminography

Almost all specialised (=LSP) dictionaries require some form of terminological analysis, with the exception of a few types of dictionaries, for example a terminological register (in the narrow sense

of the term²) or a frequency dictionary. If a dictionary is supposed to include culture-dependent terminology, such as a legal dictionary, contrastive analysis is (or at least should be) obligatorily involved. Traditionally, such studies are associated with bi- and multilingual dictionaries. However, as has been demonstrated above (cf. analysis by de Groot and van Laer), different conceptual realities can be rendered in the same natural language, leading to the compilation of monolingual contrastive terminological dictionaries.

The issue of the non-compatibility of terminological systems can equally concern ‘culture-independent’ terminology, such as the technical or scientific specialised lexis. For example, intra-lingual discrepancies may result from the co-existence of regional variants, sometimes exhibiting ‘local’ technical culture, and hence different meaning, or synonymous units of specialised vocabulary (such as professional slang or sets of ‘simplified’ terminology) used instead of, and/or alongside, terms proper in texts (discourses) of various degrees of formality and informational density (cf. highly formal professional texts vs. informal professional texts vs. popular-science texts, etc.). Similarly, the interlingual equivalence of ‘culture-independent’ terminological systems, although significantly higher than in the case of culture-dependent systems, is almost never complete. Two points need to be considered here:

1. The factors that cause conceptual discrepancies arise, *inter alia*, from the dissimilar historical development of some branches of science/technology/etc. across countries, and hence different styles of reasoning. An example to illustrate this problem is the definition of ‘outer space’ in aero-/astronautics — a notion that is necessary to nominate someone an ‘astronaut’. In the United States the adopted altitude (which a would-be astronaut has to exceed) equals 50 miles (approximately 80 kilometres), while the Fédération Aéronautique Internationale puts the figure at 100 kilometres).³ Simultaneously, the existence of idiosyncratic terminological/conceptual systems cannot be omitted from contrastive terminological considerations.
2. In the case of dynamic scientific and technical domains, new theories and developments frequently render some (older) concepts obsolete, resulting in changes in terminological systems, with some terms withdrawn from use, some new terms entering the lexicon, and — most importantly — some terms changing their definition, scope or applicability. This can lead to divergences between different systems, mainly on account of the fact that some secondary terminological systems can be slow in adopting terminological (conceptual) modifications and alterations.

As has been mentioned above, the construction of terminological dictionaries representing two or more terminological systems generally requires the analysis and comparison of the respective conceptual systems, which will then become *tertia comparationis* for further linguistic and/or conceptual analysis of the terminological sets extracted from respective specialised texts. The need for such analysis stems from the dual nature of the term: that of a linguistic sign and that of a ‘carrier’ of specialist knowledge (Lukszyn & Zmarzer, 2006, p. 23).

The application of contrastive studies to specialised dictionary-making, which one might refer to as a ‘contrastive/comparative approach to terminography’, appears at various stages in the dictionary compilation process. Several issues which arise in such studies from the perspective of practical terminography will be discussed in the following sections (3.1 and 3.2). When applied to a dedicated contrastive dictionary project, this approach gives rise to a specific structure of mono- and bilingual contrastive terminological dictionaries. The parameters of such dictionaries will be discussed in section 3.3.

²i.e. a dictionary that is composed of a mere list of terms.

³Another element which possibly evokes different frames (and/or pictures) in people’s minds are units of measurement (cf. for example: square foot vs. square metre).

3.1 General steps in specialised dictionary-making

As with general lexicographic practice, specialised dictionary-making involves a number of steps that lead to the compilation of a dictionary which meets the needs of users, is a tool of undistorted specialised knowledge transfer, and which becomes a high quality marketable commodity.

The fundamental steps include:

- designing (planning) a dictionary structure*,
- undertaking user studies (drafting a user profile, researching user needs and potential dictionary usage),
- terminographic analysis of existing secondary sources (e.g. dictionaries),
- source text selection and corpus/corpora compilation*,
- extraction of relevant terminographic data, including terms, definitions, exemplary sentences, etc.*,
- creation of a terminographic database*,
- lemma selection*,
- terminological analysis of the data*,
- writing dictionary entries,
- construction of the megastructure (putting all the data, including lemma list, indexes, bibliographies, etc. in a structure (book form or electronic format),
- final editing of the dictionary.

The following paragraphs will discuss the application of contrastive studies (or at least some elements of them) at various stages of the specialised dictionary compilation process (marked by an asterisk (*) in the list above).⁴ In some cases, the use of such analysis has already been briefly discussed in this paper (terminographic analysis and user studies), and will not be further expanded on due to the space limitations of this article.

3.2 The application of contrastive studies in specialised dictionary-making

3.2.1 Dictionary design and planning

One of the initial steps in any dictionary-making process is establishing the need to compile a dictionary for a particular group of users, i.e. a work exhibiting a specific set of parameters and presenting a specific type of linguistic (or other) data. The decisions taken at this stage will influence the compilation procedures and the final shape of the terminographic product. As has been mentioned above, almost all types of mono-, bi- and multilingual terminological dictionaries require some form of contrastive studies at both the linguistic and conceptual levels. The character of the specialised lexicon influences the depth of such analyses (consider culture-dependent *versus* culture-independent terminology). In culture-dependent dictionaries, an explicit presentation of the results of contrastive analysis comes to the forefront, either in the form of a separate encyclopaedic section, or as elements of individual entries (definitions, glosses, tags, etc.). As Bergenholtz and Tarp state: “On account of the existing cross-cultural differences, a

⁴With the exception of some predominantly ‘technical’ lexicographic activity.

comparative description of the dictionary subject matter within the respective culture area is often required too.” (Bergenholtz & Tarp, 1995, p. 61). In culture-independent works, the contrastive element is present where differences become evident and are important from the point of view of the function of the dictionary (=needs of the target users).

If a terminological dictionary is designed as a contrastive/comparative work, the contrastive element will be the focal point throughout the compilation process, and will influence the choice of terminological sources (for example, parallel texts), types of terminological data extracted (for example, definitions from multiple and varied sources, additional contexts, etc.), types of lemma included (for example, contrast-bearing terms only), analysis procedures, and elements of the mega- and microstructure (as exemplified by the model dictionary structure, see section 3.3).

3.2.2 Source text selection and corpus/corpora compilation

The primary source for any terminological dictionary is a corpus of authentic texts, primarily on account of the fact that such a corpus is a carrier of the most accurate linguistic and terminological data. Corpus tools allow for the most reliable and objective studies of a language. However, specialised (=LSP) corpora are rather scarce, and the terminographer is often left to their own devices, which in practical terms involves the compilation of one’s own corpus for terminographic purposes (Łukasik, 2014a). In fact, nowadays the compilation of a corpus can be performed fully automatically, yet such *ad hoc* corpora are of rather poor quality and lack the necessary metadata.

According to Pearson, in order to collect a reliable, i.e. a representative specialised corpus, the texts to be gathered need to be:

- (a) written (as spoken texts have to be transcribed, which may lead to the deformation of the professional message),
- (b) published (as this will validate the reliability of the material, especially if the publishing house ranks highly in terms of reputation),
- (c) produced by an acknowledged individual or institution,
- (d) factual (i.e. representing what is known to exist, or believed to exist).
(Pearson, 1998, p. 60)

However, this list should be perceived as being selective, insofar as it applies to dictionary projects that aim to present highly formal parts of the specialised lexicon. In each specific terminographic project, care must be taken to include an appropriate range of texts, i.e. one that will provide adequate coverage of the field in question (Bowker, 2003, p. 162) and deliver the required terminographic data (Łukasik, 2014a, p. 76). Clearly, terminographic projects which attempt to describe the informal part of specialised communication necessitate a very different approach to the scope and quality of texts in the corpus than the one proposed by Pearson. Such collections should include mainly spoken texts and texts published beyond the sphere of mainstream publishing.

Two types of corpora are of special interest to a lexicographer/terminographer applying contrastive methods, namely the comparable and the parallel corpus.

- The comparable corpus is one that is composed of two or more collections of texts (subcorpora) in two or more languages, with the texts sharing — to the greatest extent possible — the same features, such as topic/domain, genre, register sublanguage, or style, but without being translations of one another.
- The parallel corpus is one composed of two or more collections of texts (subcorpora), representing original texts in the source language and their translations in the target language(s). (cf. Waliński, 2005)

The applicability of both types of corpora in terminography-driven contrastive analyses will depend on the dictionary project at hand, especially as each corpus will yield different types of results. From this perspective, comparable corpora provide more accurate data on each of the languages (hence: terminological systems) analysed, and may constitute the descriptive phase for more general cross-language comparisons. Parallel corpora, meanwhile, allow for more detailed contrastive studies of terminological systems through translation equivalence (= *tertium comparationis*), as well as research into other cross-linguistic phenomena, such as terminological borrowings and terminological internationalisms. Undeniably, corpus enrichment in the form of part-of-speech or semantic annotation can greatly enhance the results obtained (for semantic annotation in parallel corpora see: Koseska-Toszewa, 2013; for the practical use of such annotated corpora see: D. Roszko & R. Roszko, 2013, Koseska-Toszewa, R. Roszko, 2016).

3.2.3 Extraction of terminographically-relevant data and the creation of a terminographic database

- (a) As mentioned above, the type of terminographic data extracted from corpora is dependent on the type of the dictionary being compiled. However, some types of data are obligatory in a terminological dictionary, in particular terms or other linguistic units to be included in the lemma list, along their definitions, if available, from the corpus. Also relevant from the perspective of contrastive studies are: illustrations of term use, collocational patterns, and grammatical and semantic categories obtained from annotated corpora, term equivalents, etc.

An important provision needs to be made with regard to the extraction of terms: if quantitative (statistics-based) methods are to be used, the results obtained in this manner should be deemed preliminary and must be paralleled with a qualitative assessment, sometimes undertaken with the help of a specialist from the field in question. This is particularly so when dealing with culture-bound terminology and the extraction of contrastively- or comparatively-relevant units.

- (b) A terminographic database is a repository of all the relevant linguistic, terminological and extralinguistic data from the perspective of a given dictionary project. In order to fully exploit a database, and possibly broaden its usefulness beyond the particular project for which it has been created, it should take the form of a so-called maximising database, i.e. one in which the database record registers as much data on the relevant headword (or topic) as is available from primary (corpus) and secondary (textbooks, other dictionaries and terminological databases, etc.) sources. In this way, it can be further developed into a universal terminographic database. The availability of universal databases could greatly enhance the quality of contrastive terminological studies.

3.2.4 Lemma selection

Lemma selection is not equivalent to term extraction from a corpus, as the former draws from the already existing terminographic database (provided one was created). The choice of headwords (a lemma list) is utterly dependent on the dictionary function (type) and user needs. A terminological dictionary can include all field-specific terms, as well as other words which are available from a corpus (and registered in the database) and significant from the perspective of the specialised domain in question. Alternatively, it can present only those terms, or other specialised lexical units, that are of special interest to the user, for example neologisms, terms with new meanings or terms exhibiting certain degrees of divergence, the latter extracted as a result of contrastive analysis of two or more terminological systems.

One specific issue requires attention when compiling some bi- and multilingual dictionaries. If such terminographic works are designed as bi- or multidirectional dictionaries, that is to say

dictionaries presenting two lemma lists, i.e. $L1 \rightarrow L2$ / $L2 \rightarrow L1$ (for bilingual works) or a greater number of such lists, i.e. for example $L1 \rightarrow L2$, $L3$ / $L2 \rightarrow L1$, $L3$ / $L3 \rightarrow L1$, $L2$ (for multilingual works), it is not possible to use a word-list reversal technique to create a lemma list of the opposite dictionary direction. This specifically concerns culture-dependent terms (Bergenholtz & Tarp, 1995, p. 101), although lack of full semantic correspondence between culture-independent terms across various languages — as has been already emphasised — renders this method only partially applicable to some specific areas and/or narrow-field scientific/technical lexis (for example, chemical nomenclature). Hence, an independent selection of lemmata (as well as their explication) needs to be undertaken in relation to each of the languages (or terminological systems) involved, while contrastive analysis must be applied in order to reveal the relations between lemmatised terms and their equivalents.

3.2.5 Terminological analysis of data

Before appearing in a dictionary, each terminological set needs to be analysed from a variety of perspectives coherent with the dictionary type, using various tools and methods. By definition, a specialised dictionary presents terms, i.e. linguistic units, that correspond to their respective concepts. Therefore, any analysis of such units requires two levels of study: linguistic and conceptual.

Linguistic analysis concerns itself with all the crucial aspects that can allow the proper identification and use of the term. The linguistic behaviour of terms can be very different from that of general language words and can vary depending on the context. Therefore, such data should be obtained from specialised, rather than general, language corpora. Among the most salient linguistic features of terms that seem to be terminographically relevant are:

- orthographic / phonetic variants,
- transliteration / transcription forms,
- morphological units (such as Greek or Latin meaningful term elements),
- part of speech characteristics,
- syntagmatic behaviour (collocation, idiom),
- style / register,
- foreign language equivalents,
- status of the term: neologism, standardised term, obsolete term, author terms, etc.

Conceptual analysis is by far the most important element of term study, since it aims to reflect the quantum of specialised knowledge carried by a term. It focuses on the reconstruction of the terminological paradigm of the term, i.e. the set of semantic or ontological relations of a given concept (represented by the term) with other concepts (terms). Among the most frequent relations analysed are hyponymy – hypernymy, meronymy – holonymy, and widely understood associative relations (Lukszyn & Zmarzer, 2006, p. 41). Interestingly enough, the traditional approach to terminology prescribes the onomasiological approach to term analysis, i.e. one in which a concept and its characteristics are to be studied before the study of language units (terms), i.e. beyond language. It is only the more recent approach that calls for the study of terms in their textual environment (corpora) (Temmerman, 2000, pp. 7–22).

Semantic relations can now be extracted automatically from specialised text corpora with an average precision of 92% (Sierra, Medina, Alarcón, & Aguilar, 2003; Acosta, Aguilar, & Sierra, 2010). In the absence of automatic tools of this kind⁵, or when the precision level is not satisfactory,

⁵Such tools are usually language-specific.

other tools such as existing ontologies, semantic networks or thesauri can be used. One example of a digitalised semantic network is WordNet, developed at Princeton University, along with corresponding national wordnets, including the Polish *Słowność*, currently being developed by the Polish CLARIN-PL language research consortium (CLARIN-PL, n.d.). The contrastive component can be obtained from (i) a juxtaposition of synsets from various wordnets or (ii) directly from a bilingual (Polish-English) component of *Słowność*, developed by way of mapping the Polish wordnet onto the Princeton Wordnet (Kędzia, Piasecki, Rudnicka, & Przybycień, 2013). Whether automatic or manual, such analyses will proceed in a semasiological-onomasiological manner, i.e. relevant data will be extracted from specialised corpora and then analysed semantically, possibly with the help of other reference works. Moreover, the cognitive value of any terminological lexicon can be provisionally assessed by quantitative measures, using predefined values ascribed to various types of terminological units (Lukszyn & Zmarzer, 2006, p. 41). Again, the importance of text corpora cannot be underestimated.

Once extracted and defined, semantic relations can help to organise the terminology presented in the dictionary into semantic fields, to develop the systematic structure of the dictionary, to build knowledge maps and, in particular, to contrast terminologies of different systems. The conceptual data can be presented in the dictionary in an explicit manner, by way of an additional semantic metalanguage (added to each entry) or a descriptive presentation in a separate dictionary appendix, or it can be presented in an implicit manner, through a system of cross-references (mediostucture).

3.3 The model of a contrastive specialised dictionary

One can envisage several types of contrastive specialist dictionaries that aim to present terminological data in a contrastive manner. The following paragraphs present a selection of metaterminographic models of dictionaries that can meet most of the criteria set for such types of reference works.

3.3.1 Monolingual dictionaries

In a monolingual terminological dictionary the contrastive constituent can pertain to either two variants of the same language (for example, American English v. British English) or two terminological systems of the same language (especially in case of culturally-dependent terminological systems, such as the legal system).

At a structural level, contrastive data can be conveyed in different parts of a monolingual contrastive terminological dictionary, particularly in its outside matter (encyclopaedic sections, knowledge maps, indexes, etc.) or in the main dictionary part (entries, cross-references, etc.).

A monolingual contrastive terminological dictionary should, as its primary function, present the primary terminological system so as to enable comparisons with the benchmark data. Such a presentation can encompass the whole system or only a certain section of it, depending on the thematic scope of the dictionary, and can take the form of a knowledge-map preceding the main body of the dictionary. In this way, any discrepancies between a concept defined in the entry and the primary terminological system can readily be identifiable, by way of cross-references to the map for example. In digital projects, such maps can form a separate database, activated on demand. In order to fully utilise the knowledge-map tool, the thesaurus-like macrostructure of the dictionary should be applied.

As some authors suggest, an important element in dictionaries presenting culturally-dependent terminology is an encyclopaedic section, or an introduction to the field defined in the dictionary (Bergenholtz & Tarp, 1995, p. 61). Alongside mainstream information, such sections could present alternative views (such as, for example dissenting opinions on legal matters), misconceptions, paradigm shifts, systemic differences, author terminology, etc. They can also reveal some linguistic features of the terminology presented.

Additionally, contrastive monolingual dictionaries can also present standard texts in the field, pointing to a specific (often highly formalised) text composition such as, for example, a sample of a statute or a patent. Such contrastive work should also focus on extralinguistic information that might influence the understanding of a specialist field, despite the common natural language used and the perceived intersubjective nature of the specialised area in question. A case in point comes from the area of mechanics, where a unit of measurement of power, known as ‘horsepower’, (a) has varying definitions, depending on the geographical region in which it is used, (b) is used according to various standards, and (c) has been employed to calculate different types of power sources (for example, electrical horsepower or boiler horsepower). The values can vary greatly as well, with mechanical horsepower roughly equal to 746 watts, and metric horsepower equivalent to approximately 736 watts (for illustration compare the definitions in: *The Illustrated Dictionary of Electronics*, 8th ed.⁶, and Larousse Online⁷).

At the microstructural level, the most important elements which introduce a contrastive dimension to a dictionary are:

- encyclopaedic and contrastive definitions,
- clear and precise sense disambiguation,
- use of references to the knowledge-map or to other entries (mediostructure),
- application of labels (or tags), indicating the terminological value, appropriateness of usage, geographical applicability, register, correctness, standardisation status, etc.,
- illustration of term use in context, collocations, standard sentences, fixed phrases,
- indication of incorrect term use and language errors.

3.4 Bi- and multilingual dictionaries

As has already been mentioned, the results of comparative studies will most naturally be presented in a bi- or multilingual dictionary, and this equally relates to terminological works. Considering the diverse nature of terminological systems, it comes as no surprise that the majority of macro- and microstructural elements that facilitate the illustration or explication of terminological data in a contrastive manner in a monolingual dictionary are also shared by their bi- or multilingual counterparts.

However, some additional requirements, and potential new ones, arise on account of the fact that more than one natural language (i.e. its lexical units) is presented in the dictionary. Since such work is aimed at presenting language elements (one- or multi-word terms, collocations, phraseological units) of two or more languages, the linguistic aspect of the terminographic data comes to the forefront. Accordingly, such linguistic phenomena as interlingual borrowings (of all kinds), terminological internationalisation, false friends, cross-language functional shifts, etc. require particular attention.

Undeniably, one of the most effective means of presenting contrastive information is through a contrastive definition. However, in a high quality terminographic works the mere fact of including a term in a dictionary confirms its terminological status⁸ (consider spelling dictionaries, for example). Similarly, a foreign language equivalent, provided that it has been chosen following

⁶**horsepower** Abbreviation, hp. A unit of power equal to 746 watts. It is generally used to express mechanical power. (Gibilisco, 2001, p. 345).

⁷**cheval-vapeur** Ancienne unité de puissance correspondant à environ 736 watts (symbole ch). Available from: <http://www.larousse.fr/encyclopedie/divers/cheval-vapeur/33221> (see: Jeuge-Maynard, 2016).

⁸The notion of ‘terminological status’ may refer to various qualities of specialised vocabulary units, and should be regarded as relative, insofar as different terminographic works may require the inclusion of lexical units of different ‘terminological status’ (cf. a standardising terminological dictionary v. a dictionary of non-standard professional lexis (≈ professional slang).

a meticulous contrastive terminological and linguistic study and has been equipped with all necessary accompanying data (disambiguation notes, usage tags, grammatical information, etc.), is worthy of a well-written definition.

One of the most salient innovations at the macrostructural level is the possibility of presenting the material in a parallel fashion, i.e. simultaneously in two or more languages. This can involve the presentation of the introductory encyclopaedic section in multiple languages. This should be done on the premise, however, that they are not necessarily translations of one another, and that they present the information that allows for a comparison of the linguistic/conceptual aspect; functionally speaking, translations of such sections could have a limiting effect. Each dictionary entry can also be presented in a parallel fashion, either in a nested technique or in a tabular form. Essentially, such entries, provided that they have been compiled in accordance with terminographic rules, are invaluable for specialists, terminologists, specialised translators, etc.

Within a dictionary microstructure, the parallel technique can manifest itself in the following forms:

- presentation of the entry term(s) and their equivalents with accompanying information essential for a full understanding of the linguistic behaviour of the units in the specialist texts of respective languages, and allowing contrastive analyses, such as:
 - sense disambiguation (for both entry terms and equivalents),
 - grammatical information/tags,
 - usage labels,
 - degree of linguistic equivalence,
 - collocational and syntactic patterns,
 - etymological information (which can greatly facilitate cross-linguistic decoding of meaning, and hence the reconstruction of respective concepts),
- comparison of concept scope and meaning by way of parallel definitions or a thesaurus part, possibly highlighting the differences between the conceptual systems (degree of conceptual equivalence),
- necessary cross-references to knowledge maps, related terms and external reading materials.

Admittedly, however, other dictionary elements are of equal importance, as they may convey additional contrastive information, such as:

- indexes of equivalents, most conceivably in the form of bilingual miniglossaries (for example, an English-Polish index in a unidirectional Polish-English dictionary),
- indexes of near synonyms, non-preferred terms, colloquialisms, internationalisms, etc.
- indexes of semantically related terms, also terms that reveal semantic shifts between two or more terminological systems presented in the dictionary,
- indexes of term elements, i.e. morphological units pertaining to complete meanings (concepts),
- grammatical tables,
- lists of false friends.

4 Conclusion

As has been demonstrated in this article, contrastive studies are inextricably linked to any form of terminographic activity, theory and practice alike. Polish metaterminography has still not achieved an optimal level of research in general, while terminographic practice, although quantitatively satisfactory, lacks in quality. The latter can be blamed partly on the reluctance of publishers to undertake thorough terminological studies prior to the compilation of any terminological dictionary, or to apply innovative methodologies which arise out of the achievements of theoretical studies, including contrastive ones.

According to a terminographic study of terminological dictionaries in English and Polish, published in Poland between 1945 and 2013 (see Łukasik, 2007, 2014b), only a few works attempt — to a greater or lesser degree — to present terminological (conceptual) data in a contrastive manner. Taking into account the structural and textual potential of traditional dictionary-making in terms of presenting linguistic and/or conceptual material in a contrastive manner, it has to be admitted that a true dictionary of this type has yet to be published. Furthermore, no electronic terminological dictionary published in Poland has fully utilised the functionalities offered by modern computer technology; needless to say, no contrastive electronic dictionary has been published to date. This state of affairs creates a strong desideratum to develop a comprehensive set of (contrastive) terminography guidelines, based on sound scientific research, that could be used by authors, editors, and publishers to address the identified need.

References

- Acosta, O., Aguilar, C., & Sierra, G. (2010). A method for extracting hyponymy-hypernymy relations from specialized corpora using genus terms. In L. A. Alemany et al. (Eds.), *Proceedings of the Workshop in Natural Language Processing and Web-based Technologies* (pp. 1–10). Córdoba: Universidad Nacional de Córdoba.
- Bergenholtz, H., & Tarp, S. (1995). *Manual of specialised lexicography*. Amsterdam: John Benjamins.
- Bowker, L. (2003). Specialised lexicography and specialised dictionaries. In P. van Starckenburg (Ed.), *A practical guide to lexicography* (pp. 154–164). Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/tlrp.6.18bow>
- CLARIN-PL. (n.d.). Retrieved 1 May 2016, from <http://clarin-pl.eu>
- de Groot, G.-R., & van Laer, C. J. P. (2006). The dubious quality of legal dictionaries. *International Journal of Legal Information*, 34(1), 64–86.
- de Groot, G.-R., & van Laer, C. J. P. (2008). *The quality of legal dictionaries: An assessment*. Maastricht: Faculty of Law.
- Gibilisco, S. (2001). *The illustrated dictionary of electronics* (8th ed.). New York: McGraw-Hill.
- Hartmann, R. R. K. (1991). Contrastive linguistics and bilingual lexicography. In F. J. Hausmann, O. Reichmann, H. E. Wiegand, & L. Zgusta (Eds.), *Dictionaries: An International Encyclopedia of Lexicography* (Vol. 3, pp. 2854–2859). Berlin: Walter de Gruyter.
- Hartmann, R. R. K. (2001). *Teaching and researching lexicography*. Harlow: Pearson Education Limited.
- Hartmann, R. R. K., & James, G. (Eds.). (2002). *Dictionary of lexicography* [DoL]. London: Routledge.
- James, C. (1980). *Contrastive analysis*. London: Longman.
- Jeuge-Maynard, I. et al. (Eds.). (2016). Larousse.fr — cheval vapeur. Retrieved 10 May 2016, from <http://www.larousse.fr/encyclopedie/divers/cheval-vapeur/33221>
- Kędzia, P., Piasecki, M., Rudnicka, E., & Przybycień, K. (2013). Automatic prompt system in the process of mapping plWordNet on Princeton WordNet. *Cognitive Studies / Études cognitives*, 2013(13), 123–141. <https://doi.org/10.11649/cs.2013.008>
- Koseska-Toszeńska, V. (2013). About certain semantic annotation in parallel corpora. *Cognitive Studies / Études cognitives*, 2013(13), 67–78. <https://doi.org/10.11649/cs.2013.004>
- Koseska-Toszeńska, V., & Roszko, R. (2016). Języki słowiańskie i litewski w korpusach równoległych Clarin-PL. *Studia z Filologii Polskiej i Słowiańskiej*, 51, 191–217. <https://doi.org/10.11649/sfps.2016.011>

- Łukasik, M. (2007). *Angielsko-polskie i polsko-angielskie słowniki specjalistyczne 1990-2006: Analiza terminograficzna*. Warszawa: Katedra Języków Specjalistycznych UW.
- Łukasik, M. (2012). Terminological dictionary as a comprehensive cognitive and linguistic tool. *Language in Different Contexts: Research Papers*, 5(1), 98–108.
- Łukasik, M. (2014a). Compiling a corpus for terminographic purposes. *Komunikacja Specjalistyczna*, 7, 71–83.
- Łukasik, M. (2014b). Najnowsze słowniki specjalistyczne z terminologią angielską i polską: Przegląd bibliograficzny w ramach analizy terminograficznej za lata 2007–2013. In M. Łukasik & B. Mikołajewska (Eds.), *Języki specjalistyczne wczoraj, dziś i jutro* (pp. 261–277). Warszawa: Wydawnictwo Naukowe IKLA. (*Studia Naukowe*, 17).
- Łukasik, M. (2016a). Perspectives of Polish-Ukrainian terminography. *The Topical Issues of Romance and Germanic Philology and Applied Linguistics*, (1–2 (11–12)), 336–343.
- Łukasik, M. (2016b). A short history of Polish-Ukrainian terminography. *Cognitive Studies / Études cognitives*, 2016(16), 65–89. <http://dx.doi.org/10.11649/cs.2016.007>
- Łukaszyn, J., & Zmarzer, W. (2006). *Teoretyczne podstawy terminologii*. Warszawa: Katedra Języków Specjalistycznych UW.
- Pearson, J. (1998). *Terms in context*. Amsterdam: John Benjamins Publishing Company.
- Pieńkos, J. (2003). *Podstawy przekładoznawstwa: Od teorii do praktyki*. Kraków: Zakamycze.
- Piotrowski, T. (1994). *Problems in bilingual lexicography*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
- Roszek, D., & Roszek, R. (2013). Experimental Polish-Lithuanian corpus with semantic annotation elements. *Cognitive Studies / Études cognitives*, 2013(13), 97–112. <https://doi.org/10.11649/cs.2013.006>
- Saloni, Z. (Ed.). (1988). *Studia z polskiej leksykografii współczesnej* (Vol. 1). Białystok: Dział Wydawnictw Filii UW.
- Sierra, G., Medina, A., Alarcón, R., & Aguilar, C. (2003). Towards the extraction of conceptual information from corpora. In D. Archer, P. Rayson, A. Wilson, & T. McEnery (Eds.), *Proceedings of the Corpus Linguistics 2003 Conference* (pp. 691–697). Lancaster: Centre for Computer Corpus Research on Language Technical Papers/ University of Lancaster.
- Temmerman, R. (2000). *Towards new ways of terminology description: The sociocognitive approach*. Amsterdam: John Benjamins Publishing Company.
- Waliński, J. (2005). Korpusy wielojęzyczne — równoległe i porównywalne. In B. Lewandowska-Tomaszczyk (Ed.), *Podstawy językoznawstwa korpusowego* (pp. 42–60). Łódź: Łódź University Press.

Acknowledgment

This work was financed by the Polish Ministry of Science and Higher Education.

The author declares that he has no competing interests.

This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 PL License (<http://creativecommons.org/licenses/by/3.0/pl/>), which permits redistribution, commercial and non-commercial, provided that the article is properly cited.