TERM BASES RELOADED

Attila IMRE[1]

Abstract

Translation industry changed enormously during the past decade due to the rapid technological advances. Cutting edge technology combined with translation practice led to the increasing importance of various databases, such as term bases and translation memories. The present article focuses on defining terms, their characteristics and how modern CAT-tools can handle term bases.

Keywords: term, term bank, definition, translation, compatibility.

Introduction

Since the advent of (semi-)automated translation, the importance of term bases has increased, as they directly contribute to consistency (quality assurance), leading to a better quality. The basic idea of term banks or term bases is rather simple: in case we have a predefined list of terms (usually in two languages), it will result in both better and faster quality.

The creation of term bases stems from terminology studies, which deals with terms and their use. Lexicology deals with words (general meaning), terminology deals with terms (specific meaning, cf. Sager, 2001c, p. 259). According to Pusztai, terminology is on the borderline of language studies, logics, ontology, informatics and sciences (1980, p. 7 cited by Á. Kis, 2005, p. 105). Thus we can say that terms are specific words or combination of words (phrases), which have a particular meaning in a special context, in particular fields. In Bowker's (2003, p. 49) definition "[t]erminology is the discipline concerned with the collection, processing, description and presentation of terms, which are lexical items belonging to specialised subject fields."

Term banks are "among the first linguistic applications of computers" (Bowker, 2003, p. 50), containing huge collections of terms in an electronic format.[2] As Bowker details, in order to ensure efficiency, they are multilingual, covering many specialised subject fields; they are also very dynamic with frequent updates, even if they cannot provide "exhaustive up-to-date coverage".

They should also be differentiated from *glossaries*, another very popular term related to specialised words. Dictionary definitions (Oxford, Cambridge) of 'glossary' usually mention that it is an alphabetical list of terms (a brief dictionary) with their definitions/explanations belonging to a specific subject often placed at the back of a book, which may be difficult to understand. It may also contain non-standard language items, such as words belonging to a dialect, but most typically it is about technical terms. The following table contains a sample from Gouadec's glossary, letter *A*:

[1] Lecturer PhD., Sapientia University, Târgu-Mureş

[2] Examples: *Eurodicautom, Termium, Normaterm, Grand dictionnaire terminologique* (Bowker, 2003, p. 50).

BDD-A2887 © 2013 Universitatea Petru Maior
Provided by Diacronia.ro for IP 216.73.217.115 (2026-11-17 07:51:30 UTC)

Active language The language into which the translator is deemed to be able to translate professionally, *i.e.* the native language or the main language when the latter has mother tongue status.
Agent/Agency A person or company acting on behalf of a buyer of translations (*i.e.* a work provider) in order to get the work done on the best possible terms. A person or company acting on behalf of a seller of translations (*i.e.* a translator) by finding the contracts the translator needs.
Assembly (and reassembly) The assembling of components, or sections, or parts.
Assumptions Everything that is necessarily true if what is being considered or stated is true.
Automatic dictionary A dictionary accessible *via* an electronic medium, offering infinitely easier and faster access than traditional paper dictionaries.
Autonomous translator A translator whose work does not require any revision.

(Gouadec, 2007, p. 373)

Characteristics and categorisation of terms

Experts in translation have described the characteristics of terms, which may serve as a guideline for creating them. According to Á. Kis, two preliminary conditions must be fulfilled from the outset:

1. preciseness (meaning is only for a given concept, no overlap whatsoever is allowed with other terms);

2. explicitness (the concept must be clearly covered by the term), excluding polysemy.

Thus we can conclude the basic requirements for terms (cf. Heltai, 2004, pp. 28–29; Á. Kis, 2005, p. 107):

- Terms have only one meaning and have no synonyms;
- Terms have a clearly defined meaning;
- Terms are always used in the same sense;
- Terms are used only by a certain group of speakers belonging to a specialty;
- The majority of terms is a compound word or a combination of words;

On the negative scale:

- The meaning of terms cannot be extended or reduced, thus they are independent from context, pragmatic factors;
- Terms are not characterised by connotation;
- Terms are not characterised by emotional meaning.

BDD-A2887 © 2013 Universitatea Petru Maior
Provided by Diacronia.ro for IP 216.73.217.115 (2026-11-17 07:51:30 UTC)

Anyway, theory and practice never fully match, which is exactly the case for terms as well. So it may happen that terms overlap, have multiple meanings and they even have emotional content (Heltai, 2004, p. 32). This is why many studies start with either defining terms, or the author explains which senses of multiple-meaning terms are used. There is no doubt that professionals, experts, scholars of various fields of study think differently of their subject of investigation when met in common, everyday day (e.g. *water* as H_2O while taking a bath cf. Heltai, 2004, p. 36), making us think that terms and context may not be separated.

Although technical terms (*terminus technicus*) belong to the language of a particular science (Á. Kis, 2005, p. 106), which come into being by definition, this is rarely possible in our age of technical revolution. Thus never before have there been so many overlapping terms, wrong terms, foreign terms (especially of English origin) than today, as we have no time for a systematic arrangement (Á. Kis, 2004, p. 47). As a result, we witness the use of "joker terms", such as *gadget* (Ro: *dispozitiv, chestie*, Hu: *kütyü*), and they may serve well (with or without code numbers) until they are commercialised and a proper name / term must be given.

Sager describes terminology compilation as a static process, consisting of "identifying, isolating and describing terminological units" (2001b, p. 251), mentioning that terminologists creating terms may use synthesis, without extending to larger textual units. In fact online access to term banks "was one of the earliest envisaged CAT tools" (Somers, 2003c, p. 20).

Terms may be categorised from the point of view of naturalisation. A new term may be temporary ("provisionally named term", cf. Sager) or 'final' ("definitive establishment of a new term-concept pair"). Virtually any method may function when a new term is coined, such as borrowing, loan, paraphrase, parallel translation, adaptation, calques or complete new creation (Sager, 2001b, p. 252), which may be influenced by the foreign language. Smaller languages usually tend to resist borrowing or loan words, but this may be true for rivalling languages as well (cf. French, German and English in Europe), hence Sager's 'permissive' languages.

There is a growing number of term banks, requiring a certain theoretical knowledge. ISO 1988 (The International Organization for Standardization) is "concerned with providing guidance on the creation of terms" (Sager, 2001b, p. 254-255) by stating a few rules:

- Precise reference;
- As economical as possible;
- Lexically systematic, conforming to morpho-phonetic rules
- Allowing word-formation (composition, derivation);
- recognizable meaning independently of any specific context.

In case there are competing terms, economy (shorter term, easier to write and remember), precision (less ambiguous) and appropriateness (more widely used) should be considered (Sager, 2001c, p. 256). As he explains, terms should not overlap, which is a

BDD-A2887 © 2013 Universitatea Petru Maior
Provided by Diacronia.ro for IP 216.73.217.115 (2026-11-17 07:51:30 UTC)

typical problem in case of dictionary entries. Thus clarity, effectiveness and unambiguous reference should always decide the proper term. If terms are long, experts will shorten it, or in case there are variants, they may bear features of social, formal, or geographical stratification. No wonder that translators often deal with variants, not always being able to 'hit the nail on the head' with parallel terms.

The difficulty lies in the fact that many terms are also present in our everyday life. English is particularly predisposed to 'raise' common words to the level of specific terms (e.g. *disk*), but other languages also do that, combined with English loan-words, especially in "international" fields, such as aviation, law or medicine (with Anglicized Latin terms). Terms deriving from computer sciences also show this tendency, as the table below illustrates:

EN	*RO*	*HU*
computer	*calculator**	*számítógép**
	*computer***	*kompjúter***
		*kalkulátor****
hard disk drive	*unitate de stocare**	*merevlemez**
	HDD, hdd	*HDD, hdd*
	hard	*hard***
	hard disk?*	*hard disk?*
	unitate de hard disk?	*winchester***
	disc rigid?	*wincsi***
		*vinyó***
* standard, ** sub-standard, *** adopted from Ro.		

Table 1. Parallel terms in Romanian and Hungarian

Synonyms (and antonyms) may constitute another area of study within terminology, as *runner-up* or *second position* is not always the same. And finally, there are cases when *gold, blood, water* may be words (common speech) or terms (chemistry, biology, hydrography / surveying).

Property and searching terms

A proper question is whether terminology belongs to theory or translators. According to Á. Kis, today this is primarily connected to translators (Á. Kis, 2005, p. 105), even if he accepts that during history this was different, as terminology is a by-product of lexicology.

The position of Á. Kis is very interesting if we compare it to Gutt's: "[a] technical translator has no right to create neologisms…whilst an advertiser or propaganda writer can use any linguistic resources he requires" (2000, p. 388), so the chance for a translator to choose an inadequate term is very high, unless (s)he 'owns' a proper term bank or (s)he

BDD-A2887 © 2013 Universitatea Petru Maior
Provided by Diacronia.ro for IP 216.73.217.115 (2026-11-17 07:51:30 UTC)

is a specialist in the field. And this brings about the question of effective search in a term base (retrieval, fuzzy match, pre-translation, term extraction, cf. Bowker, 2002, p. 90). Retrieval may include double quotation marks (exact match) or the proper use of wildcard characters (special symbols, such as *, ?, %, _, !, a. k. a. Booleans). For instance, the asterisk (*) may stand for zero, one or more characters, thus a search of *manage** has offered '*manage*', '*management*', '*manage a* public relations crisis'.

Disambiguation of concepts has been a constant problem in communication, and terminology studies aim at creating and standardizing consistent terminology for various fields. In the initial stage this may include one language, but if it is related to translation, then two or more languages are involved (bilingual or multilingual terminology).

Today, when a term bank is created/extracted, it goes without saying that it must be compatible with modern CAT-tools (translation environments). That is why Kis & Mohácsi recommend creating the term base in a Microsoft Excel file, with proper headings: *source text language, target text language,* and each cell should contain only one item (B. Kis & Mohácsi-Gorove, 2008), meaning that one source term should be inserted. Once the list is ready, we can easily convert them into a .csv file, which can be used in translation environments (CAT-tools) as a term base.

The (r)evolution of computer and translation software resulted in terminology management software as well, namely "applications designed to process the terminology required for an ongoing translation ... These are to be distinguished from full-blown database management systems that allow users to create generic or specialised terminology databases" (Gouadec, 2007, p. 270). This is in fact a specific 'term bank', which is to be used in specific text translations. Bowker clarifies that term banks and term bases are intended for human translators, so typically there is no need for additional information once the translator knows the grammar of the involved languages But she explains the 'problem' of TB from a different angle:

> In contrast, the type of information needed by a machine is very different. Machines are not intelligent and will not be able to understand definitions or contextual examples, nor will they have an innate knowledge of grammatical systems or of real-world situations. Detailed grammatical information, such as part of speech, gender, and number must be explicitly recorded in a highly structured way in machine-readable terminology resources. Morphological data, particularly for irregular plural formations or verb conjugations, will also be required by machines. (Bowker, 2003, pp. 52–53)

And we should add to the list some of our favourite insurmountable 'roadblocks' of MT: metaphors (where no grammar information will ever help), puns (on the borderline with realia).

BDD-A2887 © 2013 Universitatea Petru Maior
Provided by Diacronia.ro for IP 216.73.217.115 (2026-11-17 07:51:30 UTC)

Conclusions

Freelance translators working by themselves are often faced with multiple problems deriving from the translation of terms, and dictionaries cannot help either. The simple reason is that dictionaries usually contain "canonized" terms, leaving from bad to worse options. These include creating a new term (cf. neologisms or *hapax legomena*, disregarding Gutt's warning), offering footnotes (disrupting the text), using an ambiguous term, foreignizing (running the risk of being unintelligible), calques (may turn unprofessional or funny), or making use of the translator's competence (explanation, total change, etc.). The translator's competence may be the best choice in many new areas, although few of them have the special skill of creating either new entries or senses (Á. Kis, 2005, p. 110).

Consistency may prove how successful a term has become, and we have come across a very original definition of terminology relying on this idea. As Kis and Lengyel (2005, p. 56) define, terminology is everything that spoils the intelligibility of translation if translated inconsistently. In their view, a consistent use is easily traced during translation hence terminology of language pairs is so successful. The context refers to both the source text and the target culture and potential readers. The standardised terminology (cf. Sager), "is closely related to machine translation (MT), a field which is developing as fast as the computer generations" (Newmark, 2003, p. 64), but this already leads us to a different topic than initially envisaged.

References:

Bowker, L. (2002). *Computer-Aided Translation Technology: A Practical Introduction*. University of Ottawa Press.

Bowker, L. (2003). Chapter 4 Terminology tools for translators. In H. L. Somers (Ed.), *Computers and Translation: A Translator's Guide* (pp. 49–65). John Benjamins Publishing Company.

Gouadec, D. (2007). *Translation As a Profession*. John Benjamins Publishing.

Gutt, E.-A. (2000). Translation as interlingual interpretive use. In *The Translation Studies Reader* (pp. 376–396). Routledge.

Heltai, P. (2004). Terminus és köznyelvi szó. In *Szaknyelv és szakfordítás* (pp. 25–45). Gödöllő: Szent István Egyetem.

Kis, Á. (2004). Gyakorlati terminológia. In *Szaknyelv és szakfordítás* (pp. 46–52). Gödöllő: Szent István Egyetem.

Kis, Á. (2005). Terminusalkotás: a terminológiai helyzet és a terminológiai szerep. In *"Mindent fordítunk, és mindenki fordít" Értékek teremtése és közvetítése a nyelvészetben* (pp. 105–112). Bicske: Szak Kiadó.

Kis, B., & Lengyel, I. (2005). A fordítás számítógépes eszközei. In *"Mindent fordítunk, és mindenki fordít" Értékek teremtése és közvetítése a nyelvészetben* (pp. 53–59). Bicske: Szak Kiadó.

BDD-A2887 © 2013 Universitatea Petru Maior
Provided by Diacronia.ro for IP 216.73.217.115 (2026-11-17 07:51:30 UTC)

Kis, B., & Mohácsi-Gorove, A. (2008). *A fordító számítógépe*. Bicske: Szak Kiadó.

Pusztai, I. (1980). A bécsi terminológiai iskola elmélete és módszertana. *Magyar Nyelvőr*, *104*(1), 3–16.

Sager, J. C. (2001a). Terminology, theory. In M. Baker (Ed.), *Routledge Encyclopedia of Translation Studies* (pp. 258–262). Routledge.

Sager, J. C. (2001b). Terminology, applications. In M. Baker (Ed.), *Routledge Encyclopedia of Translation Studies* (pp. 251–255). Routledge.

Sager, J. C. (2001c). Terminology, standardization. In M. Baker (Ed.), *Routledge Encyclopedia of Translation Studies* (pp. 255–258). Routledge.

Somers, H. (2003). Chapter 2 The translator's workstation. In H. L. Somers (Ed.), *Computers and Translation: A Translator's Guide* (pp. 13–30). John Benjamins Publishing Company.

BDD-A2887 © 2013 Universitatea Petru Maior
Provided by Diacronia.ro for IP 216.73.217.115 (2026-11-17 07:51:30 UTC)