Translating Technical Texts

Document Version
Accepted author manuscript

Link to publication record in Manchester Research Explorer

Citation for published version (APA):

Published in:
Cambridge Handbook of Translation

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher’s definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.

Open Access
Translating Technical Texts

Maeve Olohan, University of Manchester

1. Technical texts and technical communication

In delimiting the scope of this chapter, I take my cue from its position within the *Handbook*, in a section on translating factual genres, alongside contributions on translating non-literary prose, legal texts, medical texts and news. Beyond the distinction made between factual and non-factual, this organisation of content may reflect the consideration of specialised text as a superordinate concept, with specific domains of specialisation as subordinate concepts, e.g., medical, legal, financial, scientific, technical texts. This conceptualisation of different domains of specialisation is commonly applied to translation too, leading to a broad understanding of technical translation as the translation of texts concerned with or related to the domain of technology. However, since technology figures in so many of today’s practices, there is substantial overlap between technical and other specialised domains; consider, among others, the sub-domains of medical or legal informatics, lawtech or fintech. We might therefore make an intuitive judgement about the extent to which a text is concerned with technical, medical or legal content and categorise accordingly. There is also a tendency to bundle technical translation with scientific translation, recognising some commonalities in topics or approaches, while also acknowledging differences (Byrne 2012; Olohan 2016).

The concept of text may also be understood in different ways, depending on the theoretical starting point (see Rogers (2015, 59–70) for an outline of possible approaches). In many commercial translation contexts, reference is made instead to ‘content’, which is no less slippery theoretically but perhaps serves to emphasise particular aspects. First, verbal and non-verbal semiotic resources may be combined in communicating technical content; consider the central role of graphics, and increasingly video, in technical communication settings. Second, such technical content may not exhibit prototypical features of linear (verbal) text; for example, strings of menu items in software interfaces or entries in product catalogues may deviate markedly from standards of textuality such as cohesion and coherence. Third, technical content may not be produced through traditional authoring practices. For example, in a practice known as topic-based authoring, technical documentation is produced by retrieving different topic-based chunks or snippets of content that have previously been written, defined (through coding of elements and content using mark-up language), stored in that standardised format and then retrieved for reuse (Lanier 2018). Thus, a user manual published as a PDF may traditionally have been authored as a linear narrative but is
now likely to be a modular assembly of numerous pieces of topic-based content, some of which could also be used in online help files, or even in a marketing brochure. Additionally, a single source of content may be used to produce a range of output types for different media or publishing channels, e.g., for print, website or mobile app. Responsive design delivers dynamic formatting so that the content can be viewed optimally, regardless of the platform on which it is published. However, like topic-based authoring, this practice of single-source authoring can have an impact on decisions about what information to communicate and how (ibid.)

Rather than consider further how a technical text might be defined and delimited, it may be more productive to focus our attention on the practices in which technical content figures. From a practice-theoretical perspective, the social realm is conceptualised as a “nexus” or “plenum” of practices (Schatzki 2001, 2; 2016, 6). Thus, to understand the nature of one practice, we can investigate both how it is constituted and how it connects with other practices. As indicated above, technical communication involves authoring practices of various kinds, and it is helpful to consider the practice of technical translation as being closely connected to those practices of technical authoring, as is done in Maylath and St. Amant’s (2019) guide to translation and localisation in technical communication. Generally speaking, before technical content is translated into target languages (TLs), it is first produced in the source language (SL). Even if assembled in modular fashion, the content that is used to produce technical specifications, product descriptions, instruction guides, user manuals, software user interfaces, help files, etc. is usually first created by technical authors.

The practices of technical authoring and technical translation are connected in various other ways, beyond their temporal sequencing. If we understand practices as organized human activities that are constituted by a range of elements (Reckwitz 2002; Shove, Pantzar, and Watson 2012), we may note that these practices of technical authoring and technical translation share some of the materials that are used, the competences that are required and the motivations that drive the practice. For example, the content management systems that are used by authors in monolingual settings are also deployed to organise multilingual, localized content. Thus, the technical know-how enabling authors and translators to work with and contribute to these systems is among a range of competences that are shared. Translators work interlingually and technical authors monolingually but both practices also require know-how that is related to technology, language and communication. Examples of shared competences can be seen in discussions of how controlled languages are used in technical communication and translation (O’Brien 2019; Musacchio 2019) or how usability testing (e.g., Quesenbery 2003; Alexander 2013) has been applied to technical translation (Byrne 2006). In addition, scholars (Risku 2004; Gneccchi et al. 2011; Vandepitte et al. 2016) address the convergence of technical communication and technical translation in professional training.

Another connection between these authoring and translating practices relates to their teleologies; they share a common end. Drawing on the concept of genre, this end can be described as producing technical content that functions as “communicative vehicles for the achievement of goals” (Swales 1990, 46). The notion of genre encapsulates the way in which a “recognizable communicative event is characterised by a set of communicative purposes” (Bhatia 1993, 13). The communicative event is structured and conventionalised, and thus is “identified and mutually understood by the members of the professional or academic community in which it regularly occurs” (ibid.). Both technical authoring and technical translation aim at producing content that will fulfil identifiable communicative purposes for their respective discourse communities. Thus, one scholarly approach to technical content is to examine it as a set of genres. We tend to refer to genres such as technical data sheets, user manuals, software interfaces, online help, standard operating procedures, among
many others, although these are more accurately considered as labels for textual formats that have become established as ways of using language for particular communicative purposes. This focus on genre is helpful in describing and explaining the challenges faced by technical translators (see, for example, Byrne 2012, Olohan 2016, Scarpa 2021 forthcoming) as they familiarise themselves with these genres and make translation choices. Thus, part of this approach is to consider closely the features of the content that is authored, translated, disseminated, read and acted upon. Through their translation decisions, translators aim to use text-internal resources to fulfil communicative purposes, thus enabling the translated text to function as part of communicative events that occur within professional communities.

A genre-focused approach also requires us to consider text-external factors, namely, to understand discursive practices as being constituent parts of professional practices (Bhatia 2008; 2016). Bhatia also highlights substantial gaps in our understanding of how genres interact and appropriate text-external resources from one another. This notion of interdiscursivity has been explored in interlingual settings (for example in Saló and Hanell’s (2014) study of how a Swedish computer scientist performs “unprecedented genres” in Swedish by drawing on prior linguistic practices in both Swedish and English) and, to a limited extent, in legal translation settings (e.g., Scott 2018). For translation studies, it would be instructive to research those professional contexts in which translated technical content is used, and to examine the ways in which translations figure in those professional practices wherever they take place, in software development and industrial manufacturing, in laboratories and research centres, in diverse installation and operation settings, and so on. This situatedness is largely unexplored in technical translation scholarship to date but offers an avenue for future research, particularly when informed by a practice-theoretical approach to translation (Olohan 2021). The next section sets the scene by highlighting the sectors of professional activity in which technical translation is prevalent.

2. Language services and technical translation

Market research companies CSA Research, Slator and Nimdzi produce regular reports on the language services sector, drawing on survey data, interviews and reports of business activities from language service providers (LSPs). The reports, when accessible, can provide useful insights into the language services landscape and the contexts for which technical content is translated today. Here we consider this global translation activity from two perspectives. In this section we first identify the domains, known as verticals, in which language services for technical content are in demand; this also provides an insight into frequently encountered professional genres. In the following section we then examine the key characteristics of how LSPs and translators operate in their sector.

The demand for translation of technical content can be gauged from The Slator 2019 Language Industry Market Report (Faes 2019), which divides the market into ten verticals. Four verticals each account for 10-13% of the market; they are technology, travel and retail, professional services, and engineering and manufacturing. Translation activities in these verticals are focused on software, user interfaces, technical documentation, technical manuals as well as marketing materials, advertising, product descriptions, websites, customer service, regulatory and legal content. Much of the translation performed here is closely connected with other practices of internationalisation and localisation, and a large proportion of the work can be considered to be technical.

Another three verticals each account for between 5-8% of the market; they are life sciences, financial services, and aerospace and defence. Much of the content translated for these verticals is also technical, comprising manuals, technical documents and clinical trial documentation, alongside legal, regulatory, research and marketing content.
The largest vertical, accounting for almost a quarter of the overall market, is the public sector. This includes the interpreting and translation that is done for international organisations, like the European Union, and governments, administrations and authorities at regional, national or local levels. Translators and interpreters in public service settings often work with content related to healthcare, military, legal or immigration matters, and some of this material addresses technical themes. For example, GLTaC (n.d.) is contracted to supply translation services to the US Army and their translations include technical content in the form of training material, field manuals and vehicle maintenance manuals. It is also not uncommon for multiple subject domains to be addressed in material translated for national or international organisations; a recent example is the European Commission Recommendation 2020/518, a non-binding legal instrument setting out a common EU approach to the use of technology and data to combat and exit from the COVID-19 crisis. In 23 language versions, this text, with an indisputable legal significance, tackles technical themes of epidemiological surveillance and monitoring, data protection, anonymisation and aggregation, the use of medical devices, and the interoperability of mobile applications.

The two final verticals presented in the Slator report are media, accounting for almost 10%, and gaming, the smallest vertical at 2.5% in 2019. In both cases, voice-over, dubbing and subtitling practices dominate, and the core fictional genres are games, films and series, discussed elsewhere in this Handbook. However, it is worth noting that some translated media content is also technical, including audiovisual content used in marketing, advertising, training and education for technical products and services.

The verticals identified by Slator are reflected to a large extent in a European survey of LSPs and individual professional translators for the same period (ELIA et al. 2019). Almost all of the activities of individual translators fit within twelve verticals, and technical content dominates in several of these. They are, in descending order of size: legal, government, other industrial, media, finance, travel, life sciences, automotive, food, consumer electronics, telecommunications and ICT (information and communications technology). The LSPs report similar levels of activity in those verticals, with small variations in significance depending on whether the LSP generated annual revenues lower or in excess of EUR 250K. The verticals of travel, food and other industrial are more important for companies with smaller revenues, while life sciences, automotive, consumer electronics and telecommunications are more important for larger LSPs.

Another way of considering which verticals are most significant for technical translation is to identify the sectors of industry or companies whose internet presence is most multilingual, since the maintenance of multilingual websites and social networks by global companies and brands usually requires translation to play a significant role. CSA Research’s Global Website Assessment Index, based on a study of over 2,800 of the world’s most prominent websites (Sargent and Lommel 2019) identified automotive, computer and electronics, and consumer goods as the top three most multilingual verticals in 2019. CSA Research used a more fine-grained categorisation of verticals (37 compared to Slator’s 10) but it is clear that there is some general agreement between each of these different approaches to identifying key verticals for translation, and these include verticals in which technical texts of various professional genres play a central role.

3. Performing technical translation

How do LSPs and translators operate in this technical content landscape? Most language services are outsourced by commissioners to LSPs. Most LSPs, in turn, outsource translation, revision, interpreting, localisation, subtitling, etc. to freelance linguists. Thus, translators of technical content generally operate as freelancers, often via the intermediation of an LSP which, in turn, manages
translation projects that are often complex and involve multiple languages. Drawing on market analysis, it is possible to distinguish between different kinds of LSP in terms of their market and strategic positioning.

The aforementioned Slator report (Faes 2019) divides the language services sector into four segments by annual revenue. The top 10% comprise the five largest LSPs of the world, called super-agencies, with revenues over USD 200m (Transperfect, Lionbridge, LanguageLine, SDL and RWS). These companies operate globally in more than one vertical. The next 12% generate over USD 25m and are labelled leader LSPs. In most cases, they have a strong regional presence. The next 5% are the challenger LSPs, generating between USD 8m and USD 25m, often focused on specific verticals or a domestic/local market. The remaining 73% comprise all other participants in the sector, including LSPs, language departments or units within organisations, and freelancers.

CSA Research organises LSPs hierarchically into five categories (DePalma 2020), based on LSPs’ abilities to adapt to changing technologies and business practices, rather than on their annual revenue. At the bottom of DePalma’s hierarchy are the “language traditionalists”. These LSPs are described as not adapting easily, making minimal use of technology and not changing their business operations much over time. The “process reengineers”, by contrast, are more technologically aware than the traditionalists and have begun reinventing themselves in response to changes in technology and business practices but are not really exploiting AI and big data. In the middle category are the “data scientists”. These LSPs have automated more of their processes than the other two groups and have integrated machine translation into their workflows.

Taking joint top position in DePalma’s hierarchy are two categories of LSP that are more specialised than the other LSPs. They are the “knowledge processing outsourcers” and the “global content strategists” and they are described as combining “a powerful blend of technical, business skills and strategic vision”. The knowledge processing outsourcers have expertise that is specialised by vertical, service or content type, and their processes are heavily driven by machine learning and technology. An example may be seen in one of RWS’s areas of activity, intellectual property (IP) services. RWS (n.d.) provides translation and other services across the IP lifecycle, from research and development to patent filing and enforcement. The final group, the global content strategists, expertly manage content creation, processing analytics and intelligence in what DePalma calls an “end-to-end information strategy”. For example, Lionbridge (n.d.) offers “end-to-end technical content creation – in any format, in any language”. This includes technical writing, training and elearning for company employees, computer-aided design and 3D modelling and media. These two categories of highly specialised global players overlap to a large extent with Slator’s super-agencies and LSP leaders.

Both categorisations show both the diversity and the fragmentation of the language services sector. The practice of translating technical content varies in line with those vastly differing revenues and operational approaches. Detailed insights into practices in operations of smaller scales are offered by several recent studies. For example, Christensen and Schjoldager (2019) report on a study of the life cycle of technical communication in a Danish medium-sized LSP, examining the relations between documentation and translation. Gonzales and Turner (2017) describe practices of technical translation, technical communication and design in the language services department of a non-profit community organization in the US. Olohan (2018) considers the convergence of scientific translation and English-language scientific editing in an international research organisation.

4. Working with textual content
We established above that a focus on communicative events and their purposes is helpful in producing effective technical content, including translations, for specific discourse communities. We also noted that some LSPs offer technical content management services that include both writing and translation. In this section we consider some key text-internal resources that are typically deployed to serve those specific communicative purposes. Scholarly contributions on technical or specialised translation are often centred on those features, which can be seen as characteristic of languages for special purposes (Krein-Kühle 2011; Byrne 2012; Rogers 2015; Olohan 2016; Scarpa 2021 forthcoming). We begin with terminology and then outline a selection of other textual and lexicogrammatical features.

Terminology usually refers to the collection of terms for a subject domain. One view of terminology (Antia et al. 2005; Cabré 2010) is based on the idea that a specialist subject domain is made up of objects with specific properties, and that those objects may be grouped into sets or classes of objects and conceptualised as units of knowledge. Those abstracted representations are known as concepts and they can be defined with reference to the characteristics that delimit them from other, related concepts. Concepts are referred to using specialised vocabulary labels known as terms. Standardisation committees prescribe what terms are preferred for sets of agreed concepts. However, the work of standardisation committees does not necessarily reflect the more complicated situation with which technical authors and translators are confronted. As Rogers (2015, 50) explains, their experience of terminology is more varied because it encompasses not only the terms used by academic experts but also the terms favoured by practitioners, hobbyists and lay users. In addition, the classic distinction between terms, conveying special meanings, and words, conveying general meanings, is often untenable in commercial contexts where an expression’s significance for business operations may determine how ‘special’ it is (ibid.: 51).

Notwithstanding debates about how terms can be defined relative to other lexical features of texts, technical translators are expected to recognise and understand the terminological references to concepts in SL content and to communicate about those concepts in the TL using terms that will be recognisable to the TL discourse community. This requires translators to deal with numerous challenges that inevitably arise through contextual variations in terminological usage. This competence or know-how is among the elements that constitute the practice of technical translation; it extends beyond familiarity with bilingual or multilingual terminology and also encompasses conceptual understanding. The next section outlines how translators manage terminological resources.

Other characteristics of languages for special purposes may be classified in various ways. For example, Halliday, when considering the difficulties of scientific writing in English, notes that terms themselves are not difficult to master, but complexity arises because of the relations that terms have to one another (Halliday 2004, 161). Writing about “interlocking definitions”, he looks at how clusters of related concepts tend to be referred to in texts and how terms are used to define other concepts, so that the reader is required to understand all of them at the same time. Technical concepts are often organised in taxonomies, and the nature of the relations that structure the taxonomy must also be understood by the reader (ibid., 165). Moving from terms to “technical grammar”, Halliday notes that special expressions may be used in technical texts where grammatical structures are different from those used in general language; he gives the examples of risk being treated as an object so that someone can ask what happens to risk, or smoking being described as something that can increase (ibid., 167).

High lexical density is another feature of scientific writing, and also relevant when considering technical content. This is the phenomenon of packing lexical items tightly into grammatical
structures, usually into noun phrases, like increasing lung cancer death rate (Halliday 2004, 168). The use of these nominal groups can introduce syntactic ambiguity, because the relation between the different elements may not be clear, e.g., is increasing lung cancer death rate about how many people die from lung cancer or about how quickly people die from lung cancer?

Nominalisation involves the replacement of one grammatical class, often a verb phrase or verbal group, by another grammatical class, in this case a noun phrase. This kind of replacement operation in general is what Halliday (2004) terms “grammatical metaphor”. The most prevalent of these shifts is the use of nominalised forms to designate processes and qualities that might otherwise be expressed by verbs or adverbs. To quote one of Halliday’s examples, fire intensity has a profound effect on smoke injection is a reconstrual of what he describes as the congruent version, which would be something like if (a) fire is intense it injects a lot more smoke or the more intense the fire, the more smoke it injects (into the atmosphere) (ibid., 28). Nominalisation produces a higher level of lexical density and an economy of linguistic forms but also reflects a reconstrual of experience. As experiences, processes or qualities are conveyed using a nominal group rather than a verbal one, the experience, process or quality is treated as a virtual entity. Linguistically, the nominalised form can serve as a subject and can be further modified. Conceptually, the virtual entity, thus abstracted, can be theorized and can participate in other relations and processes. As actors are often removed in conjunction with grammatical metaphor, an illusion of objectivity is created, alongside impressions of technicality, rationality or authority. As argued by Halliday (ibid., 128) such features may be seen as signalling the discourse of experts. At the same time, processes and procedures lose their transparency, and personal accountability is diminished.

A final difficulty of scientific English discussed by Halliday (2004, 177) is that of semantic discontinuity, where writers make semantic leaps that a reader is expected to follow; this often entails readers having to work out whether something is new information or not, and how to relates logically to another piece of information. As Halliday suggests, these leaps may not pose difficulties for specialists but may, like other features outlined, serve to exclude those who are not considered part of that elite discourse community.

Another set of concepts that facilitate analysis of those interactional dimensions of texts is proposed by Hyland (e.g., 2000; 2005) under the umbrella label of “metadiscourse”, denoting the “self-reflective expressions used to negotiate interactional meanings in a text, assisting the writer (or speaker) to express a viewpoint and engage with readers as members of a particular community” (Hyland 2005, 37). The resources used by writers to organise their text and the readers’ interaction with the information flow are grouped together as interactive resources (Hyland and Tse 2004, 157). These include transition signals, i.e., signals of topic change, or conjunctions and adverbials that signal the logical relations between ideas. They also include frame markers, i.e., references to discourse sequences, as well as endophoric markers, referring to information in other parts of the text, and evidentials, referring to sources outside of the text (i.e., citations). Also performing this interactive function are code glosses, i.e., reformulations or exemplifications. Other resources that signal writers’ attitudes to the text and to readers, showing their stance and engagement, are categorised as interactional resources (ibid.). These include hedges, boosters, attitude markers, engagement markers and self-mentions.

The use of these various resources is not merely stylistically motivated; rather they are a means by which authors seek to fulfil certain communicative purposes with certain discourse communities. Variations across discourse communities mean that metadiscursive resources may be used differently to achieve similar communicative purposes; for example, authors of research articles in pure mathematics use less hedging than authors in many humanities disciplines, where conclusions
are often offered as tentative or where alternative interpretations are invited (McGrath and Kuteeva 2012). Patterns of usage of metadiscursive resources can also vary across languages (Vold 2006; Pisanski Peterlin 2008; Mur Dueñas 2011).

Genre analysis, as established by Swales (1990; 2004) and exemplified widely in relation to the research article, as the prototypical genre of academic and scientific domains, also provides concepts that enable us to analyse how authors fulfil communicative purposes, for example through sequences of moves and sub-moves. A move is defined as a ‘discoursal or rhetorical unit that performs a coherent communicative function in a written or spoken discourse’ (Swales 2004, 228). The widespread adoption of Anglophone text-organising conventions and patterns of moves in research articles has been acknowledged but also challenged (Swales 1997; Bennett 2007). There is also some recognition of how patterns of moves reflect different epistemologies across languages (see, for example, Martin Martin 2003; Fakhri 2004; Hirano 2009) and across different academic disciplines and cultures (Samraj 2005; Stoller and Robinson 2013; Kanoksilapatham 2015).

This selection of features is not exhaustive but rather indicative of how technical content may be analysed textually, particularly in relation to communicative purpose. As exemplified in Olohan (2016) and Scarpa (2021 forthcoming), these and other textual or syntactic features can require particular attention when translating various professional scientific and technical genres.

5. Technical translation and technologies

The practice of translating is constituted not just through fulfilment of certain communicative purposes but also through the deployment of specific materials, notably the human body and a range of tools, performing as infrastructures, devices and resources (Shove 2017; Olohan 2021). One of the stark differences between small LSPs and global players in the diverse language services landscape is likely to be in how technology is integrated into practices. In this section we narrow our focus to some of the technologies that participate prominently in technical translation practice, beginning with terminology management tools and moving to translation memory and machine translation. As noted above, these may be combined with the kind of content management systems also used by clients and technical authors, or translation management systems that are predominantly deployed by LSPs and their project managers and usually involve automation of certain business processes. Translation management systems are not discussed further here (see Heinish and Iacono (2019) and Esselink (2020) for more details). Similarly, a range of specialist tools have been designed to facilitate the management of software localisation projects; these provide project management capabilities for the tasks of translators but also those of the software developers, engineers and testers whose focus is on the functionality of the localised software applications; see Roturier (2015) for details.

Terminology can be researched and managed monolingually or multilingually. The principles of terminology research and management have been set out in international standards, e.g., ISO 704 Terminology Work – Principles and methods (British Standards Institution 2010). In addition, organisations that deal with large-scale terminology management have also established standard procedures, and many examples of their terminology work are available for consultation. For example, the termbase of the Translation Bureau of Canada, Termium Plus, is freely available to consult (Government of Canada 2009). The principles by which the Bureau’s terminology work is conducted were set out in Pavel and Nolet (2001) and have been reworked into an online tutorial, available in seven languages (LinguisTech 2011). Many other institutional termbases can be searched online and sometimes also downloaded; see, for example, the European Union’s termbase, IATE (European Union n.d.).
Alongside publicly available termbases and glossaries, technical translators often have at their disposal some client-specific or project-specific terminological resources. In addition, they may conduct their own research, for example using term extraction tools to identify term candidates in the technical content they are translating and to assist in finding terminological equivalents in TL content. Typical term extraction tools that are bundled with computer-assisted translation (CAT) software or corpus query tools seek to identify terms automatically in text by a variety of means and with varying degrees of success. MemoQ is an example of CAT software whose term extraction is based somewhat crudely on the length of term candidates (in number of characters and words) and the frequency with which they occur in the text (MemoQ n.d.). Sketch Engine, by contrast, is corpus-query software that uses both statistical measures and linguistic information to identify term candidates. It compares a selected text (the focus corpus) with a reference corpus to produce lists of single-word items (keywords) and multi-word items that occur more frequently in the focus corpus than in the reference corpus (Sketch Engine 2019). Frequency is normalised (per million words) for comparison. The rules of a language-specific term grammar are then applied to reduce the lists to those items that have the linguistic forms of term candidates in that language, for example, including only nouns and noun phrases. Term candidates are then presented as lemmas, i.e., in the base form, and the translator can proceed to judge the suitability of the proposed terms.

Translators have numerous options available for the storage, management and retrieval of terminology they have researched themselves or terms that have been provided by clients or shared by other translators. While a simple bilingual glossary might easily be compiled in a spreadsheet, more specific solutions that have been designed for translation work usually offer greater functionality and usability. These applications typically integrate with translation memory applications in a local or cloud-based CAT environment and they allow terms to be both retrieved and stored during the translation process. They start as an empty termbase to which translators or others add term entries, allowing a valuable terminological resource to be built up over time. Applications vary in the extent to which they permit or require full terminological entries to be compiled for concepts. Some cloud-based CAT tools with relatively limited terminological support simply store and retrieve SL and TL terms. Other applications encourage users to compile fuller term entries, including a definition of the concept, examples of the terms in use, sources, illustrations, etc. Most applications offer sufficient flexibility to enable a translator to record the information they find most useful. The development and adoption in the 2000s of the TBX (TermBase eXchange) international format for terminology data was important in facilitating exchange of terminological data between users and applications.

The typical CAT environment has translation memory (TM) and a translation editor at its core and, increasingly, this is integrated with neural machine translation (NMT) engines via APIs. Thus, the TM software retrieves full or fuzzy matches from its database for segments of the source text that are sufficiently similar to source text segments already stored in memory. Similarity can range from 100% matches to matches of, say, 70% to 99%; the level of acceptable similarity for fuzzy match retrieval is set by the translator. The translator can then choose to leverage, i.e., re-use, the already stored translation, with or without editing, or to ignore it. Matches are retrieved at the level of the segment, usually typographically delimited as a sentence, heading, bullet point, table cell, etc., but can also be searched at sub-segmental level.

Integration of MT with the TM can take various forms. Often (and particularly in early integrations) an MT suggestion would be generated for those segments for which there were no full or fuzzy TM matches (Zaretskaya 2019). The logic is that even a fuzzy TM match is more useful to the translator than an MT suggestion, so the TM takes precedence and the MT is on where the TM can offer no
assistance. However, where NMT engines have been specialised by domain/vertical and genre and the quality of the MT suggestions is high, the MT suggestion may be given priority over anything except a full TM match; thus, a fuzzy TM match may be considered to be inferior to the MT suggestion.

How TM and MT are integrated and how translators are expected to work with the technologies vary across use cases. Generally, translators will post-edit the MT suggestions, and edit TM matches, where applicable, to produce a translation of the requisite quality. Thus, in many settings, the focus of translator activity has shifted from translating from scratch to post-editing MT suggestions (Nunes Vieira 2020). Translators working with interactive and adaptive MT systems have a slightly different role and involvement in translation production. In interactive systems (Läubli and Green 2020), the MT suggestion is changed on the basis of what the user types, while adaptive MT systems also learn from the corrections made (Daems and Macken 2019; Karimova, Simianer, and Riezler 2018). Other tools frequently used in a CAT environment, apart from terminology retrieval, include tools to perform quality assurance (QA) checks on the translation, both automatic and manual.

Research on the use of tools in professional translation practices is growing, both in laboratory settings and in translators’ workplaces. Although predominantly concerned with tools use, many of these studies involve the practice of technical translation. They address questions related to the translator’s experience from various angles, including physical and cognitive ergonomics (e.g., Ehrensberger-Dow and O’Brien 2015; Ehrensberger-Dow and Hunziker Heeb 2016; Ehrensberger-Dow and Massey 2017), user needs and expectations (e.g., Moorkens and O’Brien 2017; LeBlanc 2017) and user preferences (e.g., Cadwell et al. 2016; García-Aragón and López-Rodríguez 2017; Cadwell, O’Brien, and Teixeira 2018). These and other emerging strands of translation scholarship (Cronin 2013; 2017; O’Hagan 2016; 2020; Kenny 2017a; 2017b; Olohan 2020) are increasingly concerned with recontextualising (technical) translation in the digital world by addressing material, discursive, economic, ecological and human centred concerns. For the foreseeable future, translation studies is likely to continue to address questions pertaining to how the practice of translation “shapes and is shaped by ongoing relationships with digital technologies” (Folaron 2020, 204), ultimately leading to a better understanding of how “the fortunes of translation are bound up with the fate of technology” (Cronin 2020, 516).

References


