



Towards a multimodally oriented theory of translation: A cognitive framework for the translation of illustrated technical texts

Anne Ketola

To cite this article: Anne Ketola (2015): Towards a multimodally oriented theory of translation: A cognitive framework for the translation of illustrated technical texts, Translation Studies, DOI: [10.1080/14781700.2015.1086670](https://doi.org/10.1080/14781700.2015.1086670)

To link to this article: <http://dx.doi.org/10.1080/14781700.2015.1086670>



Published online: 15 Oct 2015.



Submit your article to this journal [↗](#)



Article views: 10



View related articles [↗](#)



View Crossmark data [↗](#)

Towards a multimodally oriented theory of translation: A cognitive framework for the translation of illustrated technical texts

Anne Ketola

School of Language, Translation and Literary Studies, University of Tampere, Finland

ABSTRACT

This article introduces a cognitively grounded theoretical framework for the translation of illustrated technical texts, situating itself at the interface of translation studies and cognitive studies of illustrated text comprehension. The article proposes that translators of illustrated technical texts process both verbal and visual information and that, consequently, their translation solutions are built on information interpreted from the combination of two different modes. The article outlines a theoretical framework supporting these propositions by examining what has been proposed about translation as cognitive activity, and comparing this with two cognitive models of illustrated text comprehension, namely the Cognitive Theory of Multimedia Learning and the Integrated Theory of Text and Picture Comprehension. The article discusses the implications of the present discussion for future research on the translation of illustrated texts and emphasizes the importance of such research.

KEYWORDS

Multimodality; technical translation; illustration; cognitive translation studies; cognitive models of multimedia learning

Introduction

Multimodality, in essence, refers to the coexistence of more than one mode – written language, spoken language, images and so on – within a given context (Gibbons 2012, 8). Multimodally oriented research posits that all modes present in a multimodal text contribute to meaning-making. It emphasizes that while the analysis of verbal language is often an important part of the investigation of multimodal texts, verbal language is always embedded within (and interpreted in relation to) a wider multimodal context (Jewitt 2009a, 2). The consideration of multimodal issues is gaining importance in all research on communication including translation studies; after all, a large number of the texts being translated today are multimodal (Hirvonen and Tiittula 2010, 1). It is then only natural that, in recent years, translation scholars have urged the traditionally language-centred discipline to respond to these challenges (e.g. O’Sullivan 2013, 6; Kaindl 2013, 266).

Translation studies has already taken the first steps towards examining multimodal texts. Pioneering research contributions on the translation of illustrated texts include Kaindl’s (2004) research into possible translation strategies for conveying humour in

the translation of comics, and Oittinen's (e.g. 1990, 2000, 2008) work discussing picture-book translation strategies that take into account both verbally and visually presented information. Previous research into the translation of illustrated technical texts, introduced in more detail below, has mainly focused on appropriate image selection and the need to adapt images during translation. More comprehensive research into the translation of illustrated technical texts is of great importance for two reasons: firstly, technical translation is a significant area of translation, constituting a major share of the translation market (Byrne 2012, 6). Secondly, different types of images are a key feature of technical documents (Byrne 2012, 26, 54; Tercedor et al. 2009, 143). Before undertaking further research, however, preliminary theoretical framing of the subject from a cognitive perspective is in order.

The present article hence sets out to explore the cognitive implications of having the visual mode involved in the translation of technical texts. Multimodality as a research issue may naturally be approached from various perspectives. The perspective adopted by the present article could be described as cognitively grounded: it examines multimodal meaning-making from the viewpoint of the cognitive processes involved in deriving meaning from various modes. The article builds on the premise that when reading an illustrated text, readers process both verbal and visual information, words and images, and form their interpretation of the multimodal text based on information provided by both of these modes. This notion has been established by investigating various types of illustrated texts, from children's picture books to scientific texts (and, logically, various types of readers from elementary school students to educated grown-ups), with methods ranging from reader interviews (Connors 2013; Youngs and Serafini 2013) and reading comprehension tests (see e.g. Mayer [2002] and Schnotz [2005], introduced below) to eye-tracking (Hegarty and Just 1993). Since translators of illustrated texts start their work as readers, this article proposes that the same argument applies for them as well. Consequently, it is proposed that their translation solutions are built on information interpreted from the combination of verbal and visual information. This, in turn, means that images may guide the way verbal text is translated.

The article provides a theoretical rationale for these claims by tentatively outlining the cognitive processes involved in illustrated text comprehension during translation, focusing on illustrated technical texts in particular. This is done by introducing two competing models of multimedia learning, namely the Cognitive Theory of Multimedia Learning (e.g. Mayer, 2002, 2005) and the Integrated Theory of Text and Picture Comprehension (e.g. Schnotz, 2005; Schnotz and Kürschner, 2008). The models have been constructed within the field of educational psychology and they describe the cognitive processes involved in illustrated text comprehension. They are based on the view of human cognition as information processing, and are therefore compatible with the majority of the theoretical frameworks adopted within cognitive studies of translation (O'Brien 2010, 2). The models are built on an extensive body of research, both theoretical and empirical, and have much to contribute to translation studies and our understanding of how the visual mode may be involved in the act of translation.

The purpose of the article is twofold. Firstly, the article seeks to contribute towards establishing a theoretical basis for multimodally oriented research in translation studies, and to extend the traditional boundaries of the discipline to include the examination of images as an object of inquiry in their own right. Secondly and most importantly, since

the framework developed in the article articulates clear implications for further research, the article strives to push forward research questions hitherto ignored, encouraging others to further analyse, test and improve on these preliminary ideas. The article will begin by introducing previous research efforts into the translation of illustrated technical texts and discussing what has been proposed about translation as cognitive activity. It will then move on to discuss how images are interpreted and how multimodal meaning-making has been conceptualized in previous research into multimodality. It will then examine how readers construct an interpretation of an illustrated text by critically reviewing and comparing the cognitive models of multimedia learning. Lastly, the article will propose one possible interpretation for how the translator of an illustrated text may negotiate meaning from an illustrated technical text in the light of these two models and propose new pathways for empirical inquiry.

Technical and cognitive translation studies: Current state of research

The definition of technical texts adopted by Franco Aixelá (2004, 32) includes “any text or text type in which there is a specific terminology belonging to a professional or academic field”. Byrne (2012, 26–28) describes the text type in a slightly more functionally oriented manner, including all task-oriented documents that seek to help a particular audience understand how to do something or how something works. When discussing *illustrated technical texts*, the present article focusses on informative, instructive texts that explain how something works and include visual information designed for the same function, examples of which would include user guides of physics textbooks. Technical texts such as technical data sheets, expert technical reports and certificates of conformity hence fall outside the research interest of the article. The article hence builds on Byrne’s definition, keeping in mind that the field-specific terminology emphasized by Franco Aixelá is indeed a key feature of technical documents (Byrne 2012, 51).

Translation involves analysing and considering a great number of factors. The need to analyse visual information in translation is acknowledged already in Christiane Nord’s (1991) model of text analysis. The model advocates the analysis of images and other non-verbal elements together with other *intratextual factors* such as the subject matter, sentence structure and lexis, concerned with the contents of the text itself, as well as *extratextual factors* relating to the function of the source text, including the sender and the sender’s intention, the target reader, and so on (Nord 1991, 36–37). In fact, Nord (1991, 110) mentions technical texts such as manuals and operating instructions as examples of illustrated texts in which images play a particularly integral role.

Despite the importance of images in technical texts, fairly little has been said about the role of images in technical translation. In an overview of audiovisual translation, Remail and Neves briefly suggest that images in technical documents produce “new constraints, possibilities and translation problems that require increasingly creative solutions” (2007, 15), though they do not elaborate on what these could be in practice. The results of a teaching project (Tercedor et al. 2009, 144, 165) which aimed to familiarize technical translation students with image analysis confirmed that images indeed open up new possibilities for translators: the study indicated that accompanying the verbal source text with (task-appropriate) images led to more creative terminological translation solutions with less source language interference.

From a theoretical perspective, the translation of illustrated technical texts has previously been studied in detail only from the perspective of appropriate image selection. Tercedor-Sánchez and Abadía-Molina (2005) discuss the need to replace the images of the source text during the translation process in order to ensure their proper transmission to a target audience, and provide general criteria for choosing images in technical and scientific texts (for discussion on similar issues within research into localization, see e.g. Hiippala [2012]; Horton [2005]). Prieto Velasco (2009, 2012), whose main interest lies in multimodal terminological databases, classifies technical and scientific texts and images based on their level of specification, explaining which types of images are most apt to represent technical and scientific concepts of varying levels of specification. In other words, these previous research efforts into illustrated technical texts have focused on the production of source texts (or database entries). The question of how these source texts are interpreted is yet to be assessed.

The issue of what happens inside the mind of the translator has been actively scrutinized over the past few decades and advanced by developments in relevant sub-disciplines of cognitive science as well as by new methods of data acquisition (for overviews, see e.g. Shreve & Angelone [2010, 2–9]; Göpferich, Jakobsen, and Mees [2009, 1–2]). Cognitive translation process research sets out to elucidate the nature of the cognitive activities involved during the translation process, usually divided into source text comprehension, transfer between the two languages, and the production of the acquired information in the target language (Englund Dimitrova 2010, 406). The present article is not concerned with the translation process as a whole, but in the processes involved in source text comprehension and their result: in Jensen's (2010, 216–217) terms, the *mental representation* constructed of the source language message.

Advances in data acquisition methods within the field of cognitive translation studies include, for example, key logging, used to examine the production processes in translation, and eye tracking, used to shed light on the source text reading and comprehension processes (Jensen 2010, 215–216); eye-tracking data is generally assumed to be indicative of the cognitive processes involved (O'Brien 2006, 186). Eye-tracking research has been used, for instance, to assess whether reading for translation differs from reading in monolingual contexts. Jakobsen and Jensen (2008) found that translators' reading time increases when they are told they will be asked to translate the text after reading it, as opposed to simply reading it for reading comprehension. This was interpreted as their reading being affected by some type of pre-translation. Similar observations were made by Jensen (2010), who concluded that, to some extent, source text comprehension and target text production are activated simultaneously when reading for translation. Further, Ruiz et al. (2008) found, albeit not with eye-tracking methods, that reading for translation involves the activation of both lexical and syntactic matches in the target language. It may hence be concluded that reading for translation is a specific type of reading activity: whether consciously or unconsciously, it involves considering tentative translation solutions.

The act of reading a text results in an interpretation of its contents. One translator's interpretation of a source text is likely to diverge, even if only slightly, from that made by any other translator. As expressed by Muñoz Martin (2010, 175–176), "each act of understanding is unique and so are, consequently, translations". In fact, Muñoz Martin emphasizes that one of the main tenets on which a cognitive theory of translation

should be based is that it is not texts but individual interpretations that we translate. In other words, the object of translation is a unique interpretation, a mental representation, prompted by the source text segment and actualized by reading comprehension processes. In what remains of the article, I will provide theoretical support for the claim that when translating a multimodal text employing verbal and visual information, these mental representations are negotiated from the combination of the two modes.

Modes making meaning

The perception of images in the cognitive apparatus may be divided into two phases: first, extracting information (shape, colour, size, depth) from the image; and second, placing an interpretation on this information (see e.g. Eysenck and Keane 2005; Anderson 1995). Parkin (2014, 28) emphasizes that visual perception is *reconstructive*; our interpretation of an image does not develop in the mind as a photograph invariably develops in the emulsion of a film. Instead, the input from the retina of the eye provides a series of cues which are inferentially processed in order to create a somewhat unique internal representation of the image contents.

The process of “reading” images is far less straightforward than reading verbal text – a process that, by necessity, proceeds in a somewhat predetermined order. Some assumptions have traditionally been made about the order in which the viewer will scan the elements of an image. It is usually assumed that elements that are made salient or eye-catching by, for instance, colours, tonal contrast or relative size are inspected first. Further, the scan path is generally considered to be affected by the human tendency of moving one’s eyes in a particular reading direction (left to right in Western cultures), centring towards the centre of the image (Foulsham and Underwood 2008), and focusing on human faces and their gaze directions (e.g. Lautenbacher 2012).

However, apart from the properties of the image, gaze is also directed by the viewer’s goals and expectations, emotions, prior knowledge and individual preferences (Boeriis and Holsanova 2012, 262). In fact, within studies of visual cognition, a growing body of research suggests that a person’s gaze is directed by cognitive information-gathering needs more than the inherent visual salience of the depicted scene (for reviews of research, see e.g. Mills et al. [2011]; Henderson et al. [2007]). In other words, viewing images is a task-oriented activity. These notions are of great importance for the current article, since they mean that different translators may attribute different meanings to the same image, and that the same translator may attribute different meanings to the same image depending on the purpose of viewing the image. Exactly the same claims were made about reading verbal texts in the previous section.

While it is challenging to comprehensively describe the meaning produced by individual modes in isolation, it is even more arduous to describe that created in the interaction of various modes. Despite the challenge, the issue has frequently been contemplated from various theoretical perspectives. The most prevalent and well established of these are based on Halliday’s (1978) systemic functional linguistics, including social semiotic approaches (building on Kress & van Leeuwen [2001]), which mainly focus on the metafunctions of modes, and the discourse analysis approach (e.g. O’Halloran 2008), which examines metafunctions in multimodal discourse at the micro-textual level. Even though the present article approaches multimodality from a cognitive perspective

instead of concentrating on the metafunctions of modes, the conceptualization of word–image interaction is here adopted as a common conceptual ground for these differing approaches.

Hull and Nelson (2005, 225), building mainly on Kress and van Leeuwen (2001, 2006), propose that the interaction of modes creates a whole new system of signification, transcending the combined contribution of individual modes. The idea is similar to that postulated by Lemke (2002, 303), who posits that the interaction of modes gives rise to genuinely new meanings and is therefore multiplicative in nature. O'Halloran (2008, 452) refers to this multiplicative interaction as *intersemiosis*, and asserts that it results in an expansion of meaning. Further, Lemke (2002, 303) has proposed that when presented together, the meanings of two modes may enhance and complement each other, offering specificity and precision beyond the capacity of either one alone. However, the meanings conveyed by individual modes in a multimodal text are not always perfectly aligned, but may also challenge or contradict each other. The relationships between modes may hence also create tensions between the aspects of meaning in a multimodal text (Jewitt 2009b, 26). The general idea proposed by these approaches therefore seems to be that when reading illustrated texts, the reader may infer meanings that are not to be found in the verbal text or the images alone: meanings that exist only in the combination of those particular words and those particular images – and for that particular reader and reading purpose, one might add. Depending on whether the messages conveyed by these two modes are aligned or not, the modes may either specify or challenge each other.

Reading illustrated technical texts

As early as 1993, Hegarty and Just conducted a study in which they monitored the eye-fixations of test subjects reading an illustrated technical text. They found that the readers frequently interrupted their reading of the verbal text to inspect the image (an average of six times per page of 140 words). Instead of looking at the image at random, the readers fixated on the parts of the image that depicted the objects they had just read about in the verbal text (Hegarty and Just 1993, 730–731). While the findings are of utmost interest for research efforts such as the present one, the study offers no theoretical explanation as to what exactly occurs in the mind of the reader when visual and verbal information are integrated in this fashion.

Studies in educational psychology have since created cognitive models of multimedia learning that discern the process of illustrated text comprehension. The article will now introduce and compare the two most influential of these, namely the Cognitive Theory of Multimedia Learning (e.g. Mayer 2002, 2005) and the Integrated Theory of Text and Picture Comprehension (Schnotz, Bannert, and Seufert 2002; Schnotz and Bannert 2003; Schnotz and Lowe 2003; Schnotz 2005; Schnotz and Kürschner 2008). Moreover, it will consider how the process of forming a mental representation of an illustrated source text would unfold in the light of these two models, and examine where the models stand in regard to empirical research findings. However, the section will begin by discussing issues of definitions. The applicability of the models of multimedia learning for the present article depends on their definitions of *multimedia* as well as their definition of *learning*; after all, one might fairly question whether theories of multimedia learning are applicable to a framework of multimodal translation.

Defining multimedia learning

As stressed by, for instance, Kress and van Leeuwen (2001, 66–67), *multimediality*, when referring purely to the analysis of the medium, and *multimodality*, when referring purely to the analysis of the semiotic resources used to make meaning, are related yet independent concepts. However, different researchers and disciplines assign different definitions to the terms (see Lauer [2009] for an extensive review). For the purposes of their research, Mayer (2005, 32) and Schnotz (2005, 50) both define a multimedia message as a communication consisting of words (printed or spoken) and images (static or dynamic) delivered in any medium. While the definition is narrower than, for instance, Kress and van Leeuwen's (2006, 177) definition of a multimodal message – according to which a multimodal message consists of any two or more modes or “semiotic codes”, be they images, spoken or written language, music, gestures and so on – it unquestionably includes the subject of the present article, namely illustrated texts.

It is generally held that translators do not need to *learn* the content of the texts they translate, in the sense of being able to retrieve and apply the information after completing the translation task (e.g. Sager 1994, 199). Yet Mayer's definition of learning describes the process as applying “cognitive strategies to incoming information in order to make sense of it” (2005, 36). The outcome of this sense-making process is either one or several mental representations (see discussion below) stored in long-term memory. Neither of the models specifies how long these representations should be retrievable from long-term memory in order for the process to count as learning. The design of the research experiments used to test these theories reflects a view that the long-term retrievability of the acquired information is not of crucial importance: whether or not the test subjects learned the information presented to them is measured by tests of transfer (i.e. being able to use the information to solve new problems) immediately after reading it (Mayer and Gallini 1990, 719; Schnotz and Bannert 2003, 151; Schnotz and Kürschner 2008, 183; Schnotz, Bannert, and Seufert 2002, 405). One could say that the models regard learning as apprehending the essential contents of something, which undoubtedly is a necessary prerequisite for successful translation activity as well.

Further, the authors behind the multimedia models do not explicitly specify which types of images their models account for. Yet the images used in their research – as well as the research they quote – are technical in nature, describing, for instance, how lightning storms develop (Mayer 2002), how geographic time differences are determined (Schnotz and Kürschner 2008), or explaining the operating principles of pulley systems (Hegarty and Just 1993), pumps and car brakes (Mayer 2002). In other words, the models are built on analyses of precisely the type of illustrated technical texts on which the article focusses; therefore the models form a plausible basis for the present theory development.

Models of illustrated text comprehension

The Cognitive Theory of Multimedia Learning (hereafter CTML) and the Integrated Theory of Text and Picture Comprehension (hereafter ITPC) share some similar assumptions about the processes that unfold during the reading comprehension of illustrated texts. Both assume that the human mind processes information in two separate channels,

the verbal and the visual, and that readers actively select, organize and integrate verbal as well as visual information in order to comprehend an illustrated text (Schnotz 2005, 57; Mayer 2005, 31). However, the models differ in some of their basic hypotheses, the most notable being the phase of integrating verbal and visual information.

Mayer's CTML postulates that the reader proceeds through an illustrated text in small segments, selecting relevant words from the verbal text and selecting relevant parts of the image. This proposition seems quite plausible if we compare it to the observations made by Hegarty and Just (1993) introduced above. Their research subjects read a segment of verbal text and then inspected the part of the image that depicted the part of the process they had just read about. CTML suggests that the relevant verbal and visual information is then organized into a verbal model in the verbal channel and a visual model in the visual channel, respectively (Mayer 2005, 38–40). CTML proposes that the two separate models are then integrated by mapping corresponding elements and relations from one model onto the other, as well as connecting them with prior knowledge. In other words, the reader combines the information extracted from the two modes by making referential links between the two, resulting in a single, integrated representation (Mayer 2005, 40). Mayer views integrated representations as effective forms of information management and retrieval. One of the principal conclusions of Mayer's research is his "Multi-media Principle", which asserts that illustrated texts may be more readily comprehended than texts with verbal information only (2002, 352–353; 2005, 31–32).

If we were to use Mayer's model of illustrated text comprehension to elucidate the translation of an illustrated technical text, how would we perceive the process of source text comprehension? First of all, the translator would read small segments of the illustrated source text, selecting parts of verbal and visual information. The translator would then form a coherent representation of the selected verbal information and a coherent representation of the selected visual information, and then integrate these with the help of prior knowledge. The input for the particular translation task would hence be a single mental representation consisting of both verbal and visual information and the task of the translator would be to formulate a target language expression describing this representation. Based on the discussion on reading for translation introduced above, one could assume that the source text comprehension process also includes instances of pre-translation: the translator considering possible translation solutions for these mental representations whilst still reading the source text.

ITPC proposes an alternative interpretation for the process of reading comprehension. Most importantly, it questions the parallelism of processing words and images as described in Mayer's model, stressing that verbal text and images are based on different sign systems and employ different principles of representation: verbal texts are descriptive representations consisting of symbols and are therefore associated with the content they represent by means of convention only. Images, on the other hand, are depictive representations consisting of iconic signs, associated with the content they represent through common structural characteristics (Schnotz and Kürschner 2008, 177). ITPC posits that when processing verbal text and images, the reader constructs internal mental representations that are also either descriptive or depictive. As these two forms of representation are fundamentally different, ITPC maintains that they cannot be mapped onto each other in order to create a single integrated model (Schnotz and Bannert 2003, 143; Schnotz and Kürschner 2008, 177).

Instead, ITPC proposes that the reader constructs multiple complementary representations both in verbal text comprehension and in image comprehension. Descriptive representations are referred to as *propositional presentations* and depictive as *mental models*. When reading verbal text, the reader constructs a descriptive propositional representation of the selected verbal text contents. This triggers the construction of a coherent depictive mental model, representing the typical visual features of what is described in the verbal text. When viewing an image, the reader constructs a depictive mental model of the selected visual information, as well as a complementing propositional representation describing the acquired information by verbal means (Schnotz and Bannert 2003, 145–146). These mental models and propositional representations are not integrated into a single representation, but interact with each other in a continuous process of what ITPC terms “mental model construction and inspection”. Each of them may be elaborated – or contested – by playing off new information from the others (Schnotz and Bannert 2003, 147; Schnotz and Kürschner 2008, 180).

According to ITPC, the translator of an illustrated technical text would again start by selecting segments of relevant visual and verbal elements. However, the number of representations formed during source text comprehension would differ from the previous model. When reading verbal text, the translator would construct both a propositional representation of the selected verbal text contents, as well as a coherent mental model depicting its typical features in visual form. In the same fashion, when retrieving visual information, the translator would build both a depictive mental model and a descriptive propositional representation verbally describing the typical features of the acquired information. The input for the translation task would consist of all these representations, which may both complement and contest each other. Again, the translator would not necessarily read the entire source text before considering translation solutions, but might consider translation solutions for one or even several of these representations as soon as they are constructed.

Evaluation of the models

A body of research in empirical psychology conducted in the late 1980s and early 1990s (e.g. Hegarty and Just 1989, 1993; Mayer 1989; Glenberg and Langston 1992) concludes that a text is easier to comprehend when it is illustrated, thus supporting Mayer’s Multimedia Principle introduced above. In these experiments, subjects reading an illustrated text comprehended and recalled the subject matter better than the control group reading verbal text alone. Such research findings have been interpreted as confirming that readers integrate the information in verbal text and images to construct an effective, integrated model of their common referent, rather than constructing separate representations of the two (Hegarty and Just 1993, 718), thus confirming Mayer’s idea.

However, all these research settings have employed images whose content is aligned with that of the presented verbal text: the images are particularly designed to support verbal text comprehension in the best possible way. While the research findings are interesting and undoubtedly helpful for purposes of textbook design, for instance, it should be emphasized that these studies observe illustrated text comprehension from one limited perspective only. As repeatedly pointed out by Schnotz and his colleagues (Schnotz and Lowe 2003, 119; Schnotz and Bannert 2003, 153; Schnotz and Kürschner 2008, 181),

CTML ignores the fact that a subject may be visualized in more than one way. Schnotz and Bannert (2003) as well as Schnotz and Kürschner (2008) propose that not all forms of illustrating a text result in its advanced comprehension. Their research findings confirm this view. Images promote text comprehension when the image is designed to support the comprehension of the verbal text in the best possible way, but they may also have a negative effect on verbal text comprehension if the subject matter is visualized less appropriately for the given verbal context (Schnotz and Bannert 2003, 153–154; Schnotz and Kürschner 2008, 187–188). It should be emphasized that this observation further supports the premise on which the present article is based, namely that the interpretation constructed from an illustrated text is based on information provided by both modes. If this were not the case, there would be no reason why the appropriateness or inappropriateness of the image for the given verbal context should affect the way in which the verbal information is interpreted.

Therefore, while CTML is perhaps slightly more straightforward than ITPC, one should question its applicability in situations where the messages conveyed by the two modes are not perfectly aligned. In particular, one may question if a reader is able to form a single, consistent interpretation of verbal and visual information in cases where the two are clearly contradictory. Even though ITPC is better suited to account for these types of texts, this model does have its own weaknesses. For instance, Reimann (2003, 250) calls for clarifications particularly in the model construction and inspection phase of the model, which, for the purposes of the current article, is of most interest. Reimann also questions the empirical testability of the somewhat complex model, asserting that novel research designs would be needed in order to empirically support it, even though refraining from suggesting what these research designs could be like. Further, in recent years, there has also been general criticism of the information-processing view of the human mind on which these models are based (see Sorden [2013] for a review). For instance, the emotions involved in illustrated text comprehension are starting to generate research interest, yet the role of affect in the process is yet to be firmly established (Sorden 2013, 21). In conclusion, it is out of the scope and ability of any single article to determine decisively which one of the models most accurately describes the cognitive processes involved. Yet, regardless of which of the models is deemed more accurate, the implications of the above discussion for translation studies remain essentially the same: when translating an illustrated technical text, the translator's interpretation of the verbal text – and, consequently, the translation solution – may be shaped by visual information.

Conclusions and implications for future research

The purpose of the present article was to discuss the cognitive implications of having images involved in the translation of illustrated technical texts. We might propose that the translator's reading comprehension process starts with constructing separate representations of the messages conveyed by the two modes. If the messages are aligned, building solid connections between the representations, each supporting the comprehension of the other, should be relatively easy. Therefore, the result could perhaps resemble a single, fused mental representation (as in CTML). If, however, the messages conveyed by the two modes are not aligned, building meaningful connections between the representations may not be possible. The comprehension process might then involve negotiating,

debating, perhaps bargaining between the information conveyed by the two modes (as in ITPC). The conclusion that emerges is that whether or not the verbal and the visual information conveyed by the source text are merged together or contested and compared, each unavoidably affects how the other is interpreted. It follows that the translation solution, too, is built on information negotiated from two different modes.

At a more particular level, the discussion presented in the article provides the following insights to inform future research:

A verbal text coupled with images might be translated differently from the same text without images, and a verbal text coupled with one set of images might be translated differently from the same text with another set of images.

Having concluded that the combination of particular words and images may give rise to meanings not to be found in either one alone, it may be postulated that adding illustrations to a verbal text can alter the translator's interpretation of the verbal text contents. Logically, different sets of illustrations would affect the interpretation in different ways.

Images might affect one translator's solutions differently from how they would affect another's. Since a viewer's interpretation of an image is affected by the viewer's individual preferences and prior knowledge, one can presume that translators interpret images in their own, slightly differing ways. The way in which a particular image might affect translation will hence depend on the translator.

Images may consolidate a translator's interpretation of a verbal source text element.

As described earlier in the article, the information conveyed by the visual mode can specify the information conveyed by the verbal mode when their messages are aligned. It might then be possible that the information provided by the image could consolidate the translator's interpretation of a possibly ambiguous element in the verbal text.

The task of the translator may be affected if visual and verbal information challenge or contradict each other.

Referring back to the studies conducted by Schnotz and Bannert (2003) and Schnotz and Kürschner (2008) introduced above, it may be assumed that if a verbal text is accompanied by images that do not support the verbal subject matter, the images are likely to have a negative effect on verbal text comprehension. Since the translator's language processing involves verbal text comprehension just like any other reader's, these circumstances may also complicate the translator's efforts to arrive at an adequate translation solution. Lamentably, research on how the combination of visual and verbal modes is comprehended offers no insights into what may happen if the information conveyed by the two modes is contradictory, but one might reasonably assume that these circumstances, too, might complicate the task of the translator.

Different types of images affect the translation process in different ways.

The above discussion has been made in the context of technical translation, building on research investigating the interpretation of technical images. Yet it has been suggested that different types of images guide the reading of accompanying verbal text to different degrees: while explanatory technical images are likely to affect the way in which accompanying verbal information is organized and integrated within the cognitive system, images that have been added to verbal text merely to decorate it may not notably guide the cognitive processing of the verbal text (Mayer 1993, 263–267) – consider, for instance, randomly adding an image of a flower bouquet to decorate a love poem. We may therefore assume that different types of images affect the translator's interpretation of the verbal text in different ways; consequently, different types of images affect the translation of verbal text in different ways. This is to say that the fact that a source text includes an image does not automatically mean that the content of the image will affect translation.

In sum, this article has proposed that when translating illustrated technical texts, translation solutions are built on information negotiated from two different modes, even though it has to be emphasized, once more, that readers themselves decide, either consciously or unconsciously, to what extent they process images, and it is therefore highly unlikely that multimodal integration of meaning would happen at every instance of translating segments of illustrated technical texts. The notion has significant implications for translation studies. Images play an essential role in the overall meaning construction of illustrated texts: they may shape the interpretation of the verbal text and affect the way in which it is translated. The discipline should hence acknowledge images as research objects in their own right. Further, the present discussion can prove to be of importance for neighbouring disciplines as well. If a translated illustrated text segment reflects an instance of how meaning may be extracted from the interaction of verbal and visual information, studying these translation solutions offers one possible tool to exemplify the nature of the semiotic space these two modes create between them. In other words, translation studies has a lot to offer to the study of multimodal meaning-making. Yet a great deal of further research is needed to develop an understanding of the complex issue. Therefore, the article aims to encourage future research to employ, test and further develop these ideas.

Funding

This research was supported by the Emil Aaltonen Foundation and the Science Foundation of the City of Tampere.

Note on contributor

Anne Ketola is a PhD candidate in the School of Language, Translation and Literary Studies at the University of Tampere, Finland, and a member of the MULTI (Multimodality in Translation and Interpreting) research group, investigating the role of multimodality within translation studies. This article is a part of her doctoral thesis in which she examines the interaction of visual and verbal information within the translation of illustrated technical texts. She is also a co-author of the forthcoming monograph *Revoicing Picturebooks* (together with Riitta Oittinen and Melissa Garavini), examining the role of visual information within picture-book translation.

References

- Anderson, John R. 1995. *Cognitive Psychology and Its Implications*. New York: W. H. Freeman.
- Boeriis, Morten, and Jana Holsanova. 2012. "Tracking Visual Segmentation: Connecting Semiotic and Cognitive Perspectives." *Visual Communication* 11 (3): 259–281. doi:10.1177/1470357212446408.
- Byrne, Jody. 2012. *Scientific and Technical Translation Explained*. Manchester: St Jerome.
- Connors, Sean P. 2013. "Weaving Multimodal Meaning in a Graphic Novel Reading Group." *Visual Communication* 12 (1): 27–53. doi:10.1177/1470357212462812.
- Englund Dimitrova, Birgitta. 2010. "Translation Process." In *Handbook of Translation Studies*, Vol. 1, edited by Yves Gambier and Luc van Doorslaer, 406–411. Amsterdam: John Benjamins.
- Eysenck, Michael W., and Mark T. Keane. 2005. *Cognitive Psychology*. Hove: Psychology Press.
- Foulsham, Tom, and Geoffrey Underwood. 2008. "What Can Saliency Models Predict about Eye Movements? Spatial and Sequential Aspects of Fixations during Encoding and Recognition." *Journal of Vision* 8 (2): 1–17. doi:10.1167/8.2.6.

- Franco Aixelá, Javier. 2004. "The Study of Technical and Scientific Translation: An Examination of its Historical Development." *Jostrans: The Journal of Specialised Translation* 1, 29–49. http://www.jostrans.org/issue01/art_aixela.php
- Gibbons, Alison. 2012. *Multimodality, Cognition, and Experimental Literature*. New York: Routledge.
- Glenberg, Arthur M., and William E. Langston. 1992. "Comprehension of Illustrated Text: Pictures Help to Build Mental Models." *Journal of Memory and Language* 31 (2): 129–151.
- Göpferich, Susanne, Arnt Lykke Jakobsen, and Inger M. Mees. 2009. Introduction to *Behind the Mind: Methods, Models and Results in Translation Process Research*, edited by Susanne Göpferich, Arnt Lykke Jakobsen, and Inger M. Mees, 1–9. Copenhagen: Samfundslitteratur.
- Halliday, Michael A. K. 1978. *Language as a Social Semiotic: The Social Interpretation of Language and Meaning*. London: Edward Arnold.
- Hegarty, Mary, and Marcel A. Just. 1989. "Understanding Machines from Text and Diagrams." In *Knowledge Acquisition from Text and Pictures*, edited by Heinz Mandl and Joel R. Levin, 171–194. Amsterdam: Elsevier.
- Hegarty, Mary, and Marcel A. Just. 1993. "Constructing Mental Models of Machines from Text and Diagrams." *Journal of Memory and Language* 32 (6): 717–742.
- Henderson, John M., James R. Brockmole, Monica S. Castelhanos, and Michael Mach. 2007. "Visual Saliency Does Not Account for Eye Movements During Visual Search." In *Eye Movements: A Window on Mind and Brain*, edited by Roger P. G. van Gompel, Martin H. Fischer, Wayne S. Murray, and Robin L. Hill, 537–562. Amsterdam: Elsevier.
- Hiiippala, Tuomo. 2012. "The Localisation of Advertising Print Media as a Multimodal Process." In *Multimodal Texts from Around the World: Cultural and Linguistic Insights*, edited by Wendy L. Bowcher, 97–122. London: Palgrave.
- Hirvonen, Maija, and Liisa Tiittula. 2010. "A Method for Analysing Multimodal Research Material: Audio Description in Focus." *MikaEL, Electronic Proceedings of the KäTu Symposium on Translation and Interpreting Studies* 4: 1–12. http://www.sktl.fi/@Bin/40698/Hirvonen%26Tiittula_MikaEL2010.pdf
- Horton, William. 2005. "Graphics: The Not Quite Universal Language." In *Usability and Internationalization of Information Technology*, edited by Nuray Aykin, 157–187. Mahwah, NJ: Erlbaum.
- Hull, Glynda A., and Mark E. Nelson. 2005. "Locating the Semiotic Power of Multimodality." *Written Communication* 22 (2): 224–261.
- Jakobsen, Arnt Lykke, and Kristian T. H. Jensen. 2008. "Eye Movement Behaviour across Four Different Types of Reading Task." In *Looking at Eyes: Tracking Studies of Reading and Translation Processing*, edited by Susanne Göpferich, Arnt Lykke Jakobsen, and Inger Mees, 103–124. Copenhagen: Samfundslitteratur.
- Jensen, Kristian T. H. 2010. "Distribution of Attention Between Source Text and Target Text During Translation." In *Cognitive Explorations of Translation: Eyes, Keys, Taps*, edited by Sharon O'Brien, 215–236. New York: Continuum International.
- Jewitt, Carey. 2009a. "Introduction: Handbook Rationale, Scope and Structure." In *The Routledge Handbook of Multimodal Analysis*, edited by Carey Jewitt, 1–13. London: Routledge.
- Jewitt, Carey. 2009b. "An Introduction to Multimodality." In *The Routledge Handbook of Multimodal Analysis*, edited by Carey Jewitt, 14–39. London: Routledge.
- Kaindl, Klaus. 2004. "Multimodality in the Translation of Humour in Comics." In *Perspectives on Multimodality*, edited by Eija Ventola, Charles Cassidy, and Martin Kaltenbacher, 173–192. Amsterdam: John Benjamins.
- Kaindl, Klaus. 2013. "Multimodality and Translation." In *The Routledge Handbook of Translation Studies*, edited by Carmen Millán and Francesca Bartrina, 257–269. London: Routledge.
- Kress, Gunther, and Theo van Leeuwen. 2001. *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Arnold.
- Kress, Gunther, and Theo van Leeuwen. 2006. *Reading Images: The Grammar of Visual Design*. 2nd ed. London: Routledge.

- Lauer, Claire. 2009. "Contending with Terms: 'Multimodal' and 'Multimedia' in the Academic and Public Spheres." *Computers and Composition* 26 (4): 225–239. doi:10.1016/j.compcom.2009.09.001.
- Lautenbacher, Olli Philippe. 2012. "From Still Pictures to Moving Pictures: Eye Tracking Text and Image." In *Eye Tracking in Audiovisual Translation*, edited by Elisa Perego, 135–155. Rome: Aracne.
- Lemke, Jay. 2002. "Travels in Hypermodality." *Visual Communication* 1 (3): 299–325. doi:10.1177/147035720200100303.
- Mayer, Richard E. 1989. "Models for Understanding." *Review of Educational Research* 59 (1): 43–64.
- Mayer, Richard E. 1993. "Illustrations That Instruct." In *Advances in Instructional Psychology*, Vol. 4, edited by Robert Glaser, 253–284. Mahwah, NJ: Erlbaum.
- Mayer, Richard E. 2002. "Using Illustrations to Promote Constructivist Learning from Science Text." In *The Psychology of Science Text Comprehension*, edited by José Otero, José A. León, and Arthur C. Graesser, 333–356. Mahwah, NJ: Erlbaum.
- Mayer, Richard E. 2005. "Cognitive Theory of Multimedia Learning." In *Cambridge Handbook of Multimedia Learning*, edited by Richard E. Mayer, 31–48. New York: Cambridge University Press.
- Mayer, Richard E., and Joan K. Gallini. 1990. "When Is an Illustration Worth Ten Thousand Words?" *Journal of Educational Psychology* 82 (4): 715–726.
- Mills, Mark, Andrew Hollingworth, Stefan Van der Stigchel, Lesa Hoffman, and Michael D. Dodd. 2011. "Examining the Influence of Task Set on Eye Movements and Fixations." *Journal of Vision* 11 (8): 1–15. doi:10.1167/11.8.17.
- Muñoz Martin, Ricardo. 2010. "On Paradigms and Cognitive Translatology." In *Translation and Cognition*, edited by Gregory M. Shreve and Erik Angelone, 169–187. Amsterdam: John Benjamins.
- Nord, Christiane. 1991. *Text Analysis in Translation: Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis*. Translated from German by Christiane Nord and Penelope Sparrow. Amsterdam: Rodopi.
- O'Brien, Sharon. 2006. "Eye-Tracking and Translation Memory Matches." *Perspectives: Studies in Translatology* 14 (3): 185–205.
- O'Brien, Sharon. 2010. Introduction to *Cognitive Explorations of Translation: Eyes, Keys, Taps*, edited by Sharon O'Brien, 1–14. New York: Continuum International.
- O'Halloran, Kay L. 2008. "Systemic Functional-Multimodal Discourse Analysis (SF-MDA): Constructing Ideational Meaning Using Language and Visual Imagery." *Visual Communication* 7 (4): 443–475. doi:10.1177/1470357208096210.
- Oittinen, Riitta. 1990. "The Dialogic Relation between Text and Illustration: A Translatological View." *TEXTconTEXT* 5: 40–52.
- Oittinen, Riitta. 2000. *Translating for Children*. New York: Garland.
- Oittinen, Riitta. 2008. "Audiences and Influences: Multisensory Translations of Picturebooks." In *Whose Story? Translating the Verbal and the Visual in Literature for Young Readers*, edited by Maria González Davies and Riitta Oittinen, 5–18. Newcastle upon Tyne: Cambridge Scholars.
- O'Sullivan, Carol. 2013. "Multimodality as Challenge and Resource for Translation." *Jostrans: The Journal of Specialised Translation* 20: 2–14. http://www.jostrans.org/issue20/art_osullivan.php
- Parkin, Alan J. 2014. *Essential Cognitive Psychology*. Hove: Psychology Press.
- Prieto Velasco, Juan Antonio. 2009. *Traducción e imagen: la información visual en textos especializados*. Granada: Ediciones Tragatanto.
- Prieto Velasco, Juan Antonio. 2012. "A Corpus-Based Approach to the Multimodal Analysis of Specialized Knowledge." *Language Resources and Evaluation* 47 (2): 399–423. doi:10.1007/s10579-012-9204-1.
- Reimann, Peter. 2003. "Multimedia Learning: Beyond Modality." *Learning and Instruction* 13 (2): 245–252.
- Remael, Aline, and Josélia Neves. 2007. "Introduction: A Tool for Social Integration? Audiovisual Translation from Different Angles." *Linguistica Antverpiensia* 6: 11–22.

- Ruiz, Carmen, Natalia Paredes, Pedro Macizo, and Maria Teresa Bajo. 2008. "Activation of Lexical and Syntactic Target Language Properties in Translation." *Acta Psychologica* 128: 490–500. doi:10.1016/j.actpsy.2007.08.004.
- Sager, Juan C. 1994. *Language Engineering and Translation: Consequences of Automation*. Amsterdam: John Benjamins.
- Schnotz, Wolfgang. 2005. "An Integrated Model of Text and Picture Comprehension." In *Cambridge Handbook of Multimedia Learning*, edited by Richard E. Mayer, 49–69. New York: Cambridge University Press.
- Schnotz, Wolfgang, and Maria Bannert. 2003. "Construction and Interference in Learning from Multiple Representation." *Learning and Instruction* 13 (2): 141–156. doi:10.1016/S0959-4752(02)00017-8.
- Schnotz, Wolfgang, and Christian Kürschner. 2008. "External and Internal Representations in the Acquisition and Use of Knowledge: Visualization Effects on Mental Model Construction." *Instructional Science* 36 (3): 175–190. doi:10.1007/s11251-007-9029-2.
- Schnotz, Wolfgang, and Richard Lowe. 2003. "Introduction: External and Internal Representations in Multimedia Learning." *Learning and Instruction* 13 (2): 117–123. doi: 10.1016/S0959-4752(02)00015-4.
- Schnotz, Wolfgang, Maria Bannert, and Tina Seufert. 2002. "Toward an Integrative View of Text and Picture Comprehension: Visualization Effects on the Construction of Mental Models." In *The Psychology of Science Text Comprehension*, edited by José Otero, José A. León, and Arthur C. Graesser, 385–416. Mahwah, NJ: Erlbaum.
- Shreve, Gregory M., and Erik Angelone. 2010. "Translation and Cognition: Recent Developments." In *Translation and Cognition*, edited by Gregory M. Shreve and Erik Angelone, 1–13. Amsterdam: John Benjamins.
- Sorden, Stephen D. 2013. "The Cognitive Theory of Multimedia Learning." In *The Handbook of Educational Theories*, edited by Beverly J. Irby, Genevieve Brown, Rafael Lara-Alecio, and Shirley Jackson, 155–168. Charlotte, NC: Information Age.
- Tercedor, Maribel, Esperanza Alarcón-Navío, Juan A. Prieto-Velasco, and Clara I. López-Rodríguez. 2009. "Images as Part of Technical Translation Courses: Implications and Applications." *Jostrans: The Journal of Specialised Translation* 11: 143–168. http://www.jostrans.org/issue11/art_tercedoretal.php
- Tercedor-Sánchez, María Isabel, and Francisco Abadía-Molina. 2005. "The Role of Images in the Translation of Technical and Scientific Texts." *Meta: Translators' Journal* 50 (4). doi:10.7202/019857ar.
- Youngs, Suzette, and Frank Serafini. 2013. "Discussing Picturebooks Across Perceptual, Structural and Ideological Perspectives." *Journal of Language and Literacy Education* 9 (1): 189–200.