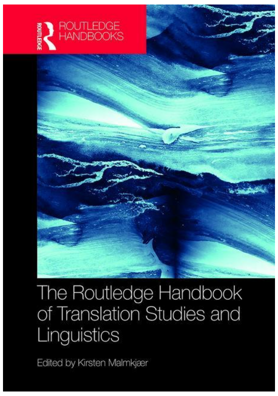

The Routledge Handbook of Translation Studies and Linguistics

Kirsten Malmkjær

Language processing in translation

Publication details
https://test.routledgehandbooks.com/doi/10.4324/9781315692845-19
Moritz Schaeffer
Published online on: 18 Dec 2017

18 Language processing in translation

Moritz Schaeffer

Introduction and definitions

There are a number of cognitive models which aim at describing linguistic processing during translation, and a large number of studies which investigate aspects of the source or target text (ST or TT). A selection of each of these kinds of studies will be reviewed here.

Within Translation Studies (TS), the use of empirical methods is relatively recent (see below). Many of the studies which investigated a particular source or target language (SL or TL) aspect and its effect on linguistic processing during translation did not integrate their findings with existing models, theories and hypotheses regarding cognitive aspects of translation; they rarely attempted to confirm or refute existing hypotheses, and focused more often on isolated phenomena. They were often aimed at testing empirical methods, and provided preliminary findings which required replication before they could be argued to apply to translation in general: if a particular aspect of a ST has an effect on linguistic processing during translation, it remains to be seen whether this effect is restricted to the two texts and languages involved in the study or whether the same kind of aspect can also be shown to have a similar effect on linguistic processing during translation of a different text involving a different language pair. One further complication makes experimental design difficult: the tools which are used to describe linguistic processing during translation need to be language independent. For example, in Danish, if a sentence starts with an adverb, the verb has to be placed before the subject while different languages (such as English or Spanish) do not require this inversion – in fact, in a declarative sentence in English, the subject is always placed before the verb. It might be more difficult to translate a declarative sentence that starts with an adverb from Danish into English than a sentence which starts with a subject, because the former requires a change in word order, while the latter does not. This is exactly what Jensen, Sjørup and Balling (2009) found. However, it might be that the difficulty observed is associated with aspects which are specific to the two languages or the two language types involved, which would mean that the resulting claims would only apply to what has been observed, and it would be difficult to extend the findings to other language combinations or to translation in general. It is therefore important to find tools which are language independent. Jensen, Sjørup and Balling (2009) and Ruiz *et al*. (2008) (for the language combination English–Spanish) used two types of sentences – critical sentences, which required a word order change when being translated; and control sentences, which could be translated using the same word order as in the source sentence. Ruiz *et al*. (2008) used Spanish sentences with a postnominal adjective (the adjective is placed after the noun) and compared this with sentences where the adjective was placed before the noun (prepositive). When translating into English, the adjective has to be placed before the noun. Participants took longer to process sentences with postnominal adjectives than prepositive adjectives. Although unlikely, it is possible that the effects observed by Ruiz *et al*. and Jensen *et al.* are related to contrastive differences between English and Spanish or Danish involving adjectives and adverbs. Schaeffer *et al*. (2016) therefore operationalised these contrastive differences by measuring the word order differences between ST and TT sentences in number of words in order to have a language-independent measure which can be correlated with associated behaviour. Schaeffer *et al*. (2016) used a measure called *Cross* which counts the word order differences between source and target sentences; i.e., if the first ST word is aligned to the fifth TT word, it has a Cross value of 5. They found that it is more difficult to translate words which have a different place in the TT sentence than in the ST sentence, and the larger these differences are, the more difficult it is to translate the words. The data on the basis of which this claim was made consisted of translations from English into Danish, Spanish, Estonian, Chinese, Hindi and German. While this is a limited number of language combinations, it shows that the effects observed apply to more than one language combination.

In sum, the challenge for Translation Process Research (TPR) is to find language-independent tools with which to describe linguistic processing during translation and to design replicable experiments in order to explain and predict behaviour during translation, so that inferences can be made concerning the cognitive processes at play during translation in general.

Historical perspectives

Much of what we know about linguistic processing during translation is based on experimental studies involving the presentation of single words. The influential study by De Groot (1992) may serve as an example. De Groot employed three translation-related tasks using single words as stimuli: "normal translation", "cued translation" and "translation recognition". In the normal translation condition, participants were presented with a single word and were asked to say out loud what this word meant in the target language. In the cued translation task, in addition to the source word participants also saw the first letter of the intended target word and had to say out loud what the target word was. Finally, in the translation recognition task, participants were shown the source and the target word and were asked to press different keys on a computer keyboard indicating whether the target word was a correct translation of the source word or not. In all three tasks, reaction times were measured, for example, the time from the moment when the source word appeared on the screen until the participant started to say the translation out loud.

In designing the stimuli, De Groot (1992) made sure that the source words either only had one translation in the TL, or had a very clearly dominant equivalent. This is important because De Groot manipulated the cognate status of the target word (how similar a target word is to its source in terms of orthography and/or phonology). Also, if there is more than one possible translation for a source word, the cue (first letter) could cause unintended confusing effects if it was not the first letter of the word the participant has in mind.

De Groot (1992) found that a number of potential characteristics of words had an effect on reaction times. These included frequency of the source and the target word (how often a word appears in a large collection of texts), imageability (the degree to which the referent of the word can evoke a mental image – it is easier to imagine *table* than *justice*), familiarity (how familiar people are with a given word – similar to frequency, though subjective), context availability (how easy it is to think of a context for a particular word), definition accuracy (how easy it is to think of a definition of a particular word) and cognate status. The more frequent, imaginable and familiar a word is, the higher its context availability, the easier it is to think of a definition and the more similar source and target words are in terms of orthography/phonology, the faster participants reacted to the stimuli. On the basis of these results, De Groot developed a model of the bilingual lexicon known as the Distributed Features Model (see Figure 18.1).

Three considerations led researchers in TS to use different methods to investigate linguistic processing during translation. Firstly, there was resistance in TS to adopting the kind of research methods De Groot was using because, it was argued, seeing a single word and saying out loud its equivalent is not what a translator normally does, so the results could not represent the cognitive processes which normally occur when a translator translates a complete text. The main and original driving force behind the development of the keylogging software Translog (see below), Arnt Lykke Jakobsen argued in 1999, was that "experiments run with Translog have ecological validity" (Jakobsen and Schou 1999, 15). Within Translog, participants normally translate longer texts while their keystrokes (and often also eye movements) are being recorded. Ecological validity relates to how representative the stimuli in an experiment are of the environment in which the mechanism under study naturally occurs. Muñoz Martín (2010, 181) argues, referring to Neisser (1976, 1987), that "results from non-ecological testing are doubtful in science", a statement that illustrates the importance that representativeness of the experimental setting has acquired in TS. Secondly, it was argued that

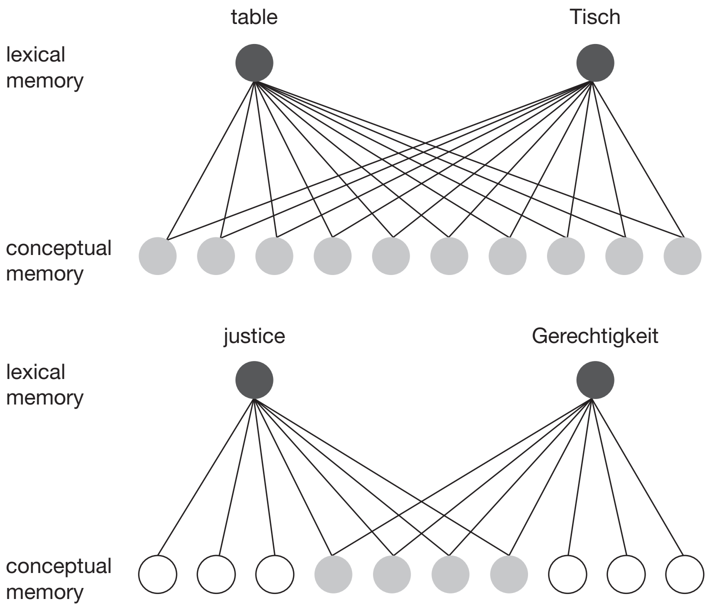

Figure 18.1 The Distributed Features Model

a bilingual who happens to have some knowledge of a second language does not engage in the same processes that a professional translator engages in when translating. The approach by researchers such as De Groot (1992, 1002) was that "word translation on its own is a useful tool to study bilingual representation and processing". De Groot (1992) therefore used bilinguals with no formal training or experience in translation. In TS, the emphasis was clearly on studying professionals or experts. Sirén and Hakkarainen (2002), echoing Krings (1986, 2001, 72) before them, argued that professional or expert translators should be studied because in that way "the pursuit of knowledge of the translation process in general, and of a successful translation process in particular, gains additional weight" (Sirén and Hakkarainen 2002, 71). In other words, the study of expert translators should be the focus given that, presumably, their processes are more successful than those of non-expert bilinguals and more relevant for translator education and theoretical insights into the translation process more generally. Thirdly, experience told TS researchers that under normal conditions, if a given text is translated by, say, 20 translators, these 20 translations are likely to differ in a number of ways. While translations for single words might be relatively stable, as soon as context plays a role, variation in the product of the process is to be expected as, e.g., Campbell (2000) and Dragsted (2012) have shown. Variability in the target texts when they are longer was one of the considerations which led to the development of Translog.

The development of Translog (Jakobsen and Schou 1999; Carl 2012) had arguably the largest effect on TPR and was, in many ways, the driving force behind what is known today as TPR. Initially, Translog could only record which keystrokes are made when, but later, eye movements and keystrokes were integrated and aligned in time and in terms of equivalence. This alignment makes it possible to, e.g., investigate a particular TT aspect and its effect on ST reading times. A number of small-scale experiments were carried out, many of which were published in the Copenhagen Studies in Language series (Göpferich, Jakobsen and Mees 2008; Alves, Göpferich and Mees 2009; Jakobsen, Mees and Göpferich 2009; Göpferich, Alves and Mees 2010; Sharp *et al*. 2011).

The development of Translog also occurred in response to a method which had been used extensively earlier (e.g., Krings 1986). The use of Think Aloud Protocols (TAPs) was the standard method to tap into cognitive processes during translation before Translog was available. During TAPs, participants are asked to verbalise their thoughts while engaged in an activity – in this case translation. The protocols are then labelled and analysed. However, contrary to those who argued that concurrent verbalisation does not interfere with the actual task at hand (Ericsson and Simon 1984), Jakobsen (2003) found that when participants translated with TAPs, they took significantly longer than when they translated without TAPs. Participants also produced the TT in significantly smaller segments, as defined by pauses of 5 seconds between keystrokes, i.e., a segment of coherent keystrokes is defined as occurring between pauses of at least 5 seconds (pauses within a segment are shorter). The fact that it took participants longer to translate the texts, and the effect on segmentation, suggest that concurrent verbalisation interferes not only with the task, but also with the associated cognitive processes. Translog was an answer to these methodological problems.

One further development had a large impact on how research into cognitive processing during translation is done, namely the development of the Translation Process Research Database (TPR-DB) (Carl, Schaeffer and Bangalore 2016), publically available under a creative commons licence (https://sites.google.com/site/centretranslationinnovation/). This database is being continually enlarged and the features it offers are constantly being refined. In 2016, the TPR-DB contained data from more than 300 different translators and more than 500 hours of text production (translation, post-editing of machine-translated text, monolingual editing of machine-translated text, monolingual copying, and more). These data were gathered in more than 1,400 sessions and contain over 600,000 translated words in 10 different TLs. Since the 2016 publication by Carl *et al.*, new target languages (Japanese) and tasks (translation with speech recognition) have been added (Carl, Aizawa and Yamada 2016). It is the only resource of its kind and offers the potential to answer a large number of questions. A number of problems which hampered earlier research are resolved with the TPR-DB: the number of participants had sometimes been rather limited and often, the lack of significant effects was attributed to the small number of participants. This is not a problem in the TPR-DB. Participants in the database all translated longer texts, unlike in the single word studies referred to earlier. Effects found in one language combination can easily be tested in a different language combination and this, together with the large number of participants, means that generalisability is less of a problem. Given that Translog is minimally invasive, it is likely that participants engage in processes they would normally engage in. There is a large and growing number of studies using the TPR-DB (https://sites.google.com/site/centretranslationinnovation/tpr-db-publications), some of which will be discussed below.

Core issues and topics

The ultimate aim of TPR is to produce a model of the cognitive processes which can explain and predict behaviour during translation. It can therefore be argued that one core issue is the role of the bilingual lexicon during translation, given that it plays a central role in bilingual language processing. Most researchers (De Groot and Starreveld 2015) now accept that language processing in bilinguals is generally non-selective: when a bilingual perceives, e.g., a word in one language, representations related to its counterpart are also activated. Possibly the most convincing study which supports this is by Wu and Thierry (2012). In this study, Chinese–English bilinguals had to press different keys on a response box depending on whether the images on the screen were squares or circles. Participants were told to ignore the English words which appeared on the screen. Unbeknown to the participants, some of the English words, when translated into Chinese, were homophones – in Chinese, they sounded like, but did not mean *circle* or *square*. Wu and Thierry recorded electrophysiological data and found an N200 effect (see below): when participants perceived the English homophones there was no effect on their behaviour – they did not press any keys; however, around 200 milliseconds after perceiving the homophones, the ERP (Event Related Potential) amplitude was larger for homophones than when participants perceived words where the translation of which was not a homophone and hence unrelated to *circle* or *square.* In other words, the brain reacted more strongly to homophones than to non-homophonic words. This kind of effect, which typically occurs after around 200 milliseconds and is called the N200 effect, is normally associated with processes involving cognitive control, i.e., participants automatically activated the Chinese equivalent and suppressed this activation immediately. Studies of bilingualism therefore focus on the mechanisms which allow multilinguals to keep their languages apart when the goal of this processing is to produce or comprehend text in one language only (Kroll *et al*. 2015).

The role of co-activation during translation was investigated by Macizo and Bajo (2004, 2006) and by Ruiz *et al*. (2008). Participants in these studies translated single sentences, and reaction times per word (during ST reading) were measured. In particular the later studies (Macizo and Bajo 2006; Ruiz *et al*. 2008) manipulated aspects of the TT which were shown to have an effect on reaction times. Bajo and colleagues aimed at testing an early and rudimentary model of the translation process proposed by Seleskovitch (1976), who argued that translation is normally carried out sequentially in that the first step is ST comprehension and only when this is complete and only once the source material is "deverbalised" can reformulation in the TL begin. Opposed to this sequential view is the assumption that representations specific to the TL are activated at the same time as SL representations are activated (horizontally and in parallel). All three studies found clear evidence against Seleskovitch's model: the evidence in these studies suggests that the TL is activated during ST reading. This might not be surprising now, though, given the overwhelming evidence there is for co-activation of the two linguistic systems in bilinguals.

Seleskovitch's early model, which was mainly designed for didactic purposes, is only one among many other attempts to produce a model of the cognitive processes that take place during translation (see Carl and Schaeffer 2017). Another such model is Halverson's (2003). This model finds its empirical validation partly in the research into the differences between translated text and original text in the same language. The focus here has been on what are called translation universals. This kind of research uses text corpora (large collections of text). Several universals have emerged from corpus-based research, such as normalisation or conventionalisation (the tendency to very frequently use linguistic patterns which are typical for the TL) (Baker 1993; but see also Malmkjær 2011). The unique items hypothesis (Tirkkonen-Condit 2004), for instance, posits that TL items which do not have a clear and obvious equivalent in the source language (SL) are less likely to be used in translated text, because "there is nothing in the source text that would trigger them off as immediate equivalents" (Tirkkonen-Condit 2004, 183). Tirkkonen-Condit (2004) found that unique items are less frequent in translated text as compared to non-translated text.

Halverson (2003) developed a model which can account for many of the findings in these corpus-based studies in terms of linguistic processing. Halverson refers to cognitive grammar (Langacker 1987) and De Groot's (1992) DFM when articulating the gravitational pull hypothesis. According to the DFM, the semantic representations of lexical items in two languages overlap to varying degrees, and words with a high degree of overlap are easier to translate. Halverson (2003) postulates that the degree of overlap has two effects: the greater the overlap, the more entrenched and the more cognitively salient these items are. Unique items in the sense described by Tirkkonen-Condit (2004) are therefore predicted to be less entrenched and less cognitively salient than words which are not unique to the TL.

In line with Langacker (1987), Halverson (2003) argues that the grammar and the lexicon form networks and the more often particular representations are used the more entrenched they become, forming proto-types. The relation between related categories of representations are also organised hierarchically, in so-called schemas. Halverson (2003, 209) explains that the networks are "characterized by global and/or local centres of gravity or prominence (prototypes and high-level schemas) that originate in various ways and that have numerous linguistic effects". Halverson further argues that during translation, "highly salient structures will exert a gravitational pull, resulting in an overrepresentation in translation of the specific TL lexical and grammatical structures that correspond to those salient nodes and configurations in the schematic network" (2003, 218). The predictions of the gravitational pull hypothesis explain, e.g., the finding from corpus-based research that translated text is more conventional than originally produced text: the salience and entrenchment of particular representations or networks makes them more likely to be used in translation than those which are less salient and entrenched. In sum, Halverson says that various translation universals "designate essentially the same thing, and represent the effects of gravitational pull exerted by category prototypes" (2003, 221). Of course, more or less implicitly underlying this theory is the assumption that SL items are co-activated with their associated TL items. However, corpus-based research has no way of making claims about the time course of the observed effects, given that the only data they use are the original and the final TT.

A study which used the TPR-DB (Schaeffer *et al*. 2016) was designed to test a model proposed by Schaeffer and Carl (2013). This model posited both early (parallel) and late (sequential) processes. Schaeffer and Carl (2013) argued that early automatic processes activate semantic and syntactic representations which are shared by the SL and the TL and that, later, monolingual vertical processes monitor the output from the early processes. Shared syntactic representations are defined in terms of the shared syntax account (Hartsuiker, Pickering and Veltkamp 2004), and shared semantic representations are defined in terms of the DFM. The shared syntax account posits that each lemma of a bilingual's languages is connected to what Hartsuiker, Pickering and Veltkamp (2004) call category nodes (such as *verb*). In addition, they are connected to combinatorial nodes (such as *active* or *passive*) and conceptual nodes (which are not language specific). Lemmas are further connected to language tags in order to allow for selective activation. However, when e.g., the Spanish word *golpear* [hit] is activated in a Spanish passive sentence, the combinatorial nodes for verb and passive are activated and hence make it more likely that subsequent language use in the other language of the bilingual will also use a passive verb – if the other language has similar syntactic structures. In sum, what the model posits is that syntax is shared across languages. The hypothesis in Schaeffer and Carl's model was that "shared representations are accessed very early during the process" (Schaeffer and Carl 2013, 174) and that during the early stages "there is no conscious control over how source and target are aligned cognitively" (Schaeffer and Carl 2013, 173).

An early, automatic effect of the TL on ST reading would be provided by an effect of, e.g., word order differences on first fixation durations. During reading, the eyes move from word to word and remain relatively stationary for certain periods of time. These stationary periods are called fixations. A first fixation duration is the time a reader spends looking at a word before either re-fixating the current word or before fixating a different word. The average fixation during normal reading is around 250ms and a great deal happens in this short period of time (Rayner 1998, 2009; Reichle, Warren and McConnell 2009), much of it thought to occur relatively automatically. The metric termed *Cross* (Carl, Schaeffer and Bangalore 2016, 26) has already been mentioned above. It denotes the word order differences between the ST and the TT segments. If the ST segment has exactly the same word order as the TT segment, then all words have a Cross value of 1. If, however, the first ST word is aligned to the sixth TT word, then the Cross value is 6. If the sixth ST word is aligned to the first TT word, then the Cross value is –6. As mentioned earlier, this feature of the TPR-DB is language independent and can therefore be used across languages. One other feature of the TPR-DB is called word translation entropy (HTra). It is computed by counting how many different TT items, which are aligned to the same ST item, there are in a corpus of a number of translations of the same ST. On the basis of the probabilities of each of these TT realisations, the distribution of these probabilities is calculated (Carl, Schaeffer and Bangalore 2016, 31). The more different TT realisations there are in a given sample of translations of the same ST, the larger the HTra value and the more likely it is that the overlap in terms of semantics is smaller. Schaeffer *et al*. (2016) found that both HTra and syntactic distortion (Cross) had a significant positive effect on first fixation durations and total reading time. Total reading time is the sum of all fixations on a particular word – irrespective of when they occur. The effect of Cross and HTra on first fixation durations probably represents early, automatic cognitive alignment, which is less effortful in the case of ST items for which the overlap between ST and TT representations in terms of syntax and semantics is greater (low HTra and Cross values). The data for the study by Schaeffer *et al*. (2016) consisted of 42,211 English ST words translated into six different target languages. While the large number of languages and the sizeable amount of data encourage confidence in the results, it should be stressed that a non-negligible amount of variation could not be explained with the predictors in the model presented by Schaeffer *et al*. (2016). In other words, while the model could make predictions with a certain degree of confidence, a possibly large number of variables which impact eye movements during translation remain unknown.

Main research methods

The principal research methods in TPR are eye tracking and keylogging. However, researchers are starting to use brain imaging techniques such as electroencephalography (EEG) and functional Magnetic Resonance Imaging (fMRI). The use of these imaging techniques is promising, although it is in its early stages within TS. One other fundamental distinction needs to be made: there are corpus-based eye movement or keylogging studies and those which tightly control stimuli and manipulate a (small number of) variable(s). Eye movement or keylogging corpus studies such as the one discussed above by Schaeffer *et al*. (2016) or the one by Balling and Carl (2014) normally control variables statistically, while experimental studies such as those by Bajo and colleagues (Macizo and Bajo 2004, 2006; Ruiz *et al*. 2008), the one by Wu and Thierry (2012) or the one by De Groot (1992) control variables in the design of the stimuli and the task. This is an important distinction and some advantages and disadvantages have already been discussed. In a typical experimental study, participants carry out one task and the stimuli consist of filler, critical and control items. Critical and control items are identical in many ways apart from one aspect or a small number of aspects. This makes it possible to make relatively strong claims regarding the effect of the manipulated variable on the dependent variable (reaction times, eye movement or keylogging measures, or ERP waveforms). Particular care is taken to make sure that the critical stimuli do not differ in any way apart from in the independent variable(s) of interest. The disadvantage of this kind of research is that, because stimuli and tasks are so highly controlled, and because sample sizes are often small, laborious replication is required to extend the results to other language combinations and participant samples – in addition to the fact that some tasks are unnatural: if, e.g., single sentences are presented one at a time in an eye tracking study, participants can normally not re-read earlier sentences and effects of the larger context cannot be measured. The dangers of corpus-based research are that a very large number of variables can have an effect on the dependent variable – many of which may be unknown. It may be that the hypothesised effect is in the data, but that it is buried under many other factors which are not included in the statistical model and which obscure the hypothesised effect. It may also be that the hypothesised effect is found, but that one or a number of unknown factors underlie the observed effect and the claim made on the basis of the observed effect should actually be attributed to (a) different (number of) aspect(s) of the data. Translation Studies is not alone in this conflict: there are a number of studies which investigate eye movements during normal reading on the basis of a corpus, such as the studies by Kennedy and Pynte (2005) or, more recently, by Cop, Drieghe and Duyck (2015). In the Cop *et al.* study, 19 bilinguals and 14 monolinguals read a whole novel while their eye movements were recorded. There are a very large number of experimental studies which investigate the effects of a particular, and highly controlled, aspect on eye movements (Rayner 1998, 2009). Ideally, findings from controlled experimental studies are tested in more natural settings and vice versa. The findings from single word studies such as the one by De Groot (1992) have been tested in studies involving single sentences (e.g., Schwartz and Kroll 2006) and in reading of a long text (e.g., Cop *et al*. 2015).

The dependent variables in eye tracking used in TPR have been imported from Psychology. The kind of research from which TPR has imported these eye movement measures was mainly interested in the early processes during reading. However, possibly the largest and most obvious difference between normal reading and reading while translating is that translation takes much longer than normal reading. Translation of course involves writing (in a different language) while normal reading does not. Kliegl *et al*. (2004) report a mean total reading time per word during reading for comprehension of 245ms (SD = 48). However, an examination of the data in the TPR-DB shows that during (monolingual) copying the mean total reading time per word on the source text is 797ms (SD = 1068), while during translation the mean total reading time per ST word is 1577ms (SD = 5824). Both latter tasks involve reading and producing a text and, on average, participants spend twice as long reading an ST word when they are translating it than when they are copying it and six times as long as when they are reading for comprehension. The co-ordination of reading and writing in addition to having to manage the demands of interlinguistic reformulation therefore requires a different set of eye movement measures than those that are currently used in Psychology and TPR. Dragsted (2010) suggests one such measure, the eye–key span (EKS) which measures the time between first fixation of a word and the first keystroke which contributes to the translation of the equivalent target word(s). This is similar to the ear–voice span in simultaneous interpreting. Schaeffer and Carl (Forthcoming) propose to use the probability with which the ST is read while the TT is being typed. Schaeffer and Carl (2013) argued that this could be used as an indicator of the degree to which processes are automated, and Schaeffer and Carl (Forthcoming) find evidence to support this claim. However, given the strong effect of the task (translation) on the late eye movement measures such as total reading time, and given that traditional eye movement studies offer very few measures which can adequately describe these late processes, TPR is in need of more suitable tools to describe eye movements on a whole text while the subject is typing the translation.

The way keylogging contributes to TPR has traditionally been by either giving access to interim solutions of a translator (e.g., Tirkkonen-Condit, Mäkisalo and Immonen 2008) or by making use of the time delay between individual keystrokes, i.e., pauses in the flow of typing (e.g., Immonen 2006). Interim solutions consist of text which is typed and then deleted. This kind of data are not accessible on the basis of the final TT, but keylogging allows the analysis of these deleted keystrokes. There are a number of features in the TPR-DB, which are unique to it, unlike the traditional eye movement measures imported from Psychology. Two will be described in detail.

The *InEff* feature (Carl, Schaeffer and Bangalore 2016, 26) describes the inefficiency with which a translation is produced. It is calculated by dividing the number of keystrokes which contributed to the production of a translation by the number of characters in the final TT. If the translator did not revise the translation in any way, *InEff* is 1; but if, say, the word was completely rewritten once, *InEff* is 2 – the higher this value, the more editing went into the translation.

The *Munit* feature (Carl, Schaeffer and Bangalore 2016, 25) indicates how many Micro Units (Alves and Vale 2011) a translator needs in order to produce a particular item. A Micro Unit is defined as a continuous typing activity between pauses (no typing) of 1 second or more. The information in the TPR-DB for a *Munit* is its duration, the duration of the typing pause that preceded it, the number of fixations that occurred (on the ST and the TT) during this *Munit*, and the time that the participant spent reading the ST while typing the TT. All durations are in milliseconds, so it is possible to analyse typing and eye movement behaviour in a fairly detailed manner. The rich features that the TPR-DB offers make it an ideal resource for the modelling of the cognitive processes during translation.

Current debates

Possibly the most well-studied model of the bilingual lexicon is the Revised Hierarchical Model (RHM) (Kroll and Stewart 1994). There are other, more recent models (Dong, Gui and MacWhinney 2005), in addition to the DFM (De Groot 1992) referred to earlier. Studies in TPR rarely base predictions on these models or attempt to test the predictions made by these models. This is partly because these models focus on lexical equivalence and are typically based on single word studies. However, García (2015, 21) argues that

> Accessing lexical equivalents of source-text words is just one of multiple mental activities during translation, and a most basic one at that. However, the basic nature of this skill does not render it trivial. On the contrary, it underscores the importance of studying lexical equivalence to understand more complex translation processes, since they will necessarily imply such a skill.

Given that the RHM makes very clear predictions regarding the impact of directionality on linguistic processing during translation – predictions which have been extensively tested – and given that basic factors such as cognate status, concreteness and L2 proficiency have been found to have reliable effects on behaviour during translation, it seems odd that these factors very rarely find their way into the design of experiments in TPR. It seems that, because of the lack of ecological validity (natural tasks and stimuli) in the many single word studies, researchers in TPR have rejected these findings. However, the reverse is also true: authors of single word studies do not tend to refer to studies which employ more natural texts and tasks – presumably because these researchers distrust results from studies which lack the kind of control employed in single word studies. Neutralising the tension between controlled experimental studies and studies which use more ecologically valid designs, and cross-fertilisation between findings from the two camps, promises to yield interesting findings and could lead to solid, generalisable results which will find their way into pedagogical and commercial applications.

Future directions

Three research trends which have already been highlighted above are likely to play an important role in future TPR. The findings from corpus-based research, i.e., studies of large corpora of translated texts (e.g., Hansen-Schirra, Neumann and Steiner 2012) could relatively easily be tested in TPR, but this has not happened to date. The model by Halverson (2003) may provide a good framework for the generation of hypotheses. Secondly, the incorporation of models and the testing of hypotheses generated on the basis of single word studies in the context of bilingualism research has not happened either and is likely to lead to interesting findings. Thirdly, the testing of findings from corpus-based TPR in controlled experiments is likely to lead to more solid models. In addition, machine learning has been used with TPR data (e.g., Martinez-Gomez *et al*. 2014). Martinez-Gomez *et al.* automatically classify process data (eye movement and keystroke records) in terms of whether and to what degree participants are professionals or not. Martinez-Gomez *et al.* were able to tell – with a reasonable degree of confidence – whether a participant was an expert or not and how many years' experience a translator had (with a margin of error of 4.15 years). This approach is likely to be exploited in future studies, because it has the potential to be useful in a number of applications (see below).

Implications for Practice

The challenge for TPR is to produce results which are relevant for the translation-related industries and the teaching of translation. So far, the findings from TPR have not achieved this aim. In the context of the post-editing of machine translation, there has been research which is more likely to be relevant for industry (e.g., Doherty, O'Brien and Carl 2010; Lacruz, Denkowski and Lavie 2014; Vieira 2014; O'Brien 2005, 2006, 2007). However, most of these studies investigate the relationship between cognitive effort and the quality of the machine-translated text – linguistic processing is subsumed under the generic term "cognitive effort" without necessarily specifying the actual processes which take place.

Further Reading

Carl, M., Bangalore, S. and Schaeffer, M. J. 2016. *New Directions in Empirical Translation Process Research: Exploring the CRITT TPR-DB*. Berlin: Springer.

Chapter 2 in this book gives a good overview of all the features in the TPR-DB. It also contains accounts of a large number of studies – all of which use the TPR-DB in one way or another.

Ferreira, A. and Schwieter, J. W. 2015. *Psycholinguistic and Cognitive Inquiries into Translation and Interpreting*. Amsterdam: John Benjamins.

This book gives a good overview of more current studies into the cognitive aspects of translation (and interpreting).

Shreve, G. M. and Angelone, E. 2010. *Translation and Cognition*. Amsterdam: John Benjamins.

This book is a good source for earlier studies and contains accounts of both empirical and more theoretical investigations.

Related topics

Corpus linguistics, translation and interpreting; Translation, interpreting and new technologies.

References

Alves, F., Göpferich, S. and Mees, I. M., eds. 2009. *Methodology, Technology and Innovation in Translation Process Research: A Tribute to Arnt Lykke Jakobsen*. Copenhagen: Samfundslitteratur.

Alves, F. and Vale, D. C. 2011. On drafting and revision in translation: a corpus linguistics oriented analysis of translation process data. In *Parallel Corpora: Annotation, Exploitation, Evaluation*. Special issue of *Translation: Corpora, Computation, Cognition* 1(1), pp. 105–122.

Baker, M. 1993. "Corpus Linguistics and Translation Studies". In *Text and Technology: In Honour of John Sinclair*, edited by M. Baker, G. Francis and E. Tognini-Bonelli, 233–250. Amsterdam: John Benjamins.

Balling, L. W. and Carl, M. 2014. "Production Time Across Languages and Tasks: A Large-Scale Analysis Using the CRITT Translation Process Database". In *The Development of Translation Competence: Theories and Methodologies from Psycholinguistics and Cognitive Science*, edited by J. W. Schwieter and A. Ferreira, 239–268. Newcastle upon Tyne: Cambridge Scholars Publishing.

Campbell, S. 2000. “Choice Network Analaysis in Translation Research”. In *Intercultural Faultlines: Research Models in Translation Studies I – Textual and Cognitive Aspects*, edited by M. Olohan, 29–42. Manchester: St Jerome.

Carl, M. 2012. “Translog-II: A Program for Recording User Activity Data for Empirical Reading and Writing Research”. In *The Eighth International Conference on Language Resources and Evaluation, 21–27 May 2012*, 2–6. Istanbul, Turkey: Department of International Language Studies and Computational Linguistics.

Carl, M., Aizawa, A. and Yamada, M. 2016. “English-to-Japanese Translation vs. Dictation vs. Post-editing: Comparing Translation Modes in a Multilingual Setting”. In *The LREC 2016 Proceedings: Tenth International Conference on Language Resources and Evaluation*, edited by N. Calzolari, K. Choukri, T. Declerck, M. Grobelnik, B. Maegaard, J. Mariani, A. Moreno, J. Odijk and S. Piperidis, 4024–4031. Portorož: ELRA.

Carl, M. and Schaeffer, M. J. 2017. “Models of the Translation Process”. In *The Handbook of Translation Studies and Cognition*, edited by J. Schwieter and A. Ferreira, 50–70. Amsterdam: John Benjamins.

Carl, M., Schaeffer, M. J. and Bangalore, S. 2016. “The CRITT Translation Process Research Database”. In *New Directions in Empirical Translation Process Research: Exploring the CRITT TPR-DB2*, edited by M. Carl, S. Bangalore and M. Schaeffer, 13–54. Cham: Springer.

Cop, U., Drieghe, D. and Duyck, W. 2015. Eye movement patterns in natural reading: a comparison of monolingual and bilingual reading of a novel. *PLoS ONE* 10(8), pp. 1–38.

Doherty, S., O’Brien, S. and Carl, M. 2010. Eye tracking as an MT evaluation technique. *Machine Translation* 24(1), pp. 1–13.

Dong, Y., Gui, S. and MacWhinney, B. 2005. Shared and separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition* 8(3), pp. 221–238.

Dragsted, B. 2010. “Coordination of Reading and Writing Processes in Translation: An Eye on Uncharted Territory”. In *Translation and Cognition*, edited by G. M. Shreve and E. Angelone, 41–62. Amsterdam: John Benjamins.

Dragsted, B. 2012, Indicators of difficulty in translation: correlating product and process data. *Across Languages and Cultures* 13(1), pp. 81–98.

Ericsson, K. A. and Simon, H. A. 1984. *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA: MIT Press.

García, A. M. 2015. Psycholinguistic explorations of lexical translation equivalents: thirty years of research and their implications for cognitive translatology. *Translation Spaces* 4(1), pp. 9–28.

Göpferich, S., Alves, F. and Mees, I. M., eds. 2010. *New Approaches in Translation Process Research*. Copenhagen: Samfundslitteratur.

Göpferich, S., Jakobsen, A. L. and Mees, I. M., eds. 2008. *Looking at Eyes: Eye-Tracking Studies of Reading and Translation Processing*. Copenhagen: Samfundslitteratur.

De Groot, A. M. B. 1992. Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18(5), pp. 1001–1018.

De Groot, A. M. B. and Starreveld, P. A. 2015. “Parallel Language Activation in Bilinguals’ Word Production and its Modulating Factors: A Review and Computer Simulations”. In *The Cambridge Handbook of Bilingual Processing*, edited by J. Schwieter, 389–415. Cambridge: Cambridge University Press.

Halverson, S. 2003. The cognitive basis of translation universals. *Target: International Journal of Translation Studies* 2(15), pp. 197–241.

Hansen-Schirra, S., Neumann, S. and Steiner, E. 2012. *Cross-Linguistic Corpora for the Study of Translations: Insights from the Language Pair English-German*. Berlin: de Gruyter.

Hartsuiker, R. J., Pickering, M. J. and Veltkamp, E. 2004. Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science* 15(6), pp. 409–414.

Immonen, S. 2006. Translation as a writing process: pauses in translation versus monolingual text production. *Target: International Journal of Translation Studies* 18(2), pp. 313–336.

Jakobsen, A. L. 2003. "Effects of Think Aloud on Translation Speed, Revision and Segmentation". In *Triangulating Translation: Perspectives in Process Oriented Research*, edited by F. Alves, 69–95. Amsterdam: John Benjamins.

Jakobsen, A. L., Mees, I. M. and Göpferich, S., eds. 2009. *Behind the Mind: Methods, Models and Results in Translation Process Research*. Copenhagen: Samfundslitteratur.

Jakobsen, A. L. and Schou, L. 1999. "Translog Documentation". In *Probing the Process in Translation: Methods and Results*, edited by G. Hansen, 1–36. Frederiksberg: Samfundslitteratur.

Jensen, K. T. H., Sjørup, A. C. and Balling, L. W. 2009. "Effects of L1 Syntax on L2 Translation". In *Methodology, Technology and Innovation in Translation Process Research: A Tribute to Arnt Lykke Jakobsen*, edited by F. Alves, S. Göpferich and I. M. Mees, 319–336. Copenhagen: Samfundslitteratur.

Kennedy, A. and Pynte, J. 2005. Parafoveal-on-foveal effects in normal reading. *Vision Research* 45, pp. 153–168.

Kliegl, R., Grabner, E., Rolfs, M. and Engbert, R. 2004. Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology* 16(1–2), pp. 262–284.

Krings, H. P. 1986. *Was in den Köpfen von Übersetzern vorgeht: eine empirische Untersuchung zur Struktur des Übersetzungsprozesses an fortgeschrittenen Französischlernern*. Tübingen: Günter Narr Verlag.

Krings, H. P. 2001. *Repairing Texts: Empirical Investigations of Machine Translation Post-Editing Processes*. Kent, OH: The Kent State University Press.

Kroll, J., Gullifer, J., Mcclain, R. and Martin, M. C. 2015. "Selection and Control in Bilingual Comprehension and Production". In *The Cambridge Handbook of Bilingual Processing*, edited by J. W. Schwieter, 485–507. Cambridge: Cambridge University Press.

Kroll, J. F. and Stewart, E. 1994. Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 33(2), pp. 149–174.

Lacruz, I., Denkowski, M. and Lavie, A. 2014. "Cognitive Demand and Cognitive Effort in Post-Editing". In *Proceedings of the AMTA 2014 Workshop on Post-Editing Technology and Practice*, edited by S. O'Brien, M. Simard and L. Specia, 73–84. Vancouver: Canada.

Langacker, R. 1987. *Foundations of Cognitive Grammar 1*. Stanford, CA: Stanford University Press.

Macizo, P. and Bajo, M. T. 2004. When translation makes the difference: sentence processing in reading and translation. *Psicológica* 25, pp. 181–205.

Macizo, P. and Bajo, M. T. 2006. Reading for repetition and reading for translation: do they involve the same processes? *Cognition* 99(1), pp. 1–34.

Malmkjær, K. 2011. "Translation Universals". In *The Oxford Handbook of Translation Studies*, edited by K. Malmkjær and K. Windle, 83–94. Oxford: Oxford University Press.

Martinez-Gomez, P., Minicha, A., Huang, J., Carl, M. and Aizawa, A. 2014. "Recognition of Translator Expertise Using Sequences of Fixations and Keystrokes". In *ETRA '14: Proceedings of the Symposium on Eye Tracking Research and Applications*, edited by P. Qvarfordt and D. Witzner Hansen, 299–302. New York: Association for Computing Machinery.

Muñoz Martín, R. 2010. "On paradigms and cognitive translatology". In *Translation and Cognition*, edited by G. Shreve and E. Angelone, 169–187. Amsterdam: John Benjamins.

Neisser, U. 1976. *Cognition and Reality*. San Francisco: Freeman.

Neisser, U. 1987. *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*. Cambridge: Cambridge University Press.

O'Brien, S. 2005. Methodologies for measuring the correlations between post-editing effort and machine translatability. *Machine Translation* 19(1), pp. 37–58.

O'Brien, S. 2006. Pauses as indicators of cognitive effort in post-editing machine translation output. *Across Languages and Cultures* 7(1), pp. 1–21.

O'Brien, S. 2007. Eye-tracking and translation memory matches. *Perspectives* 14(3), pp. 185–205.

Rayner, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124(3), pp. 372–422.

Rayner, K. 2009. Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology* 62(8), pp. 1457–1506.

Reichle, E. D., Warren, T. and McConnell, K. 2009. Using E-Z reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin and Review* 16(1), pp. 1–21.

Ruiz, C., Paredes, N., Macizo, P. and Bajo, M. T. 2008. Activation of lexical and syntactic target language properties in translation. *Acta Psychologica* 128(3), pp. 490–500.

Schaeffer, M. J. and Carl, M. 2013. Shared representations and the translation process: a recursive model. *Translation and Interpreting Studies* 8(2), pp. 169–190.

Schaeffer, M. J. and Carl, M. Forthcoming. "Translation and Non-Translational Language Use". In *Empirical Modelling of Translation and Interpreting*, edited by S. Hansen-Schirra, O. Čulo, S. Hofmann and B. Meyer. Berlin: Language Science Press.

Schaeffer, M. J., Dragsted, B., Hvelplund, K. T. and Balling, L. W. 2016. "Word Translation Entropy: Evidence of Early Target Language Activation During Reading for Translation". In *New Directions in Empirical Translation Process Research: Exploring the CRITT TPR-DB*, edited by M. Carl, S. Bangalore and M. Schaeffer, 183–210. Berlin: Springer.

Schwartz, A. I. and Kroll, J. F. 2006. Bilingual lexical activation in sentence context. *Journal of Memory and Language* 55(2), pp. 197–212.

Seleskovitch, D. 1976. "Interpretation: A Psychological Approach to Translating". In *Translation: Applications and Research*, edited by R. W. Brislin, 92–116. New York: Gardner.

Sharp, B., Zock, M., Carl, M. and Jakobsen, A. L., eds. 2011. *Human–Machine Interaction in Translation: Proceedings of the 8th International NLPCS Workshop*. Copenhagen: Samfundslitteratur.

Sirén, S. and Hakkarainen, K. 2002. Expertise in translation. *Across Languages and Cultures* 3(1), pp. 71–82.

Tirkkonen-Condit, S. 2004. "Unique Items: Over- or Under-Represented in Translated Language?" In *Translation Universals: Do they Exist?*, edited by A. Mauranen and P. Kujamäki, 177–184. Amsterdam: John Benjamins.

Tirkkonen-Condit, S., Mäkisalo, J. and Immonen, S. 2008. The translation process: interplay between literal rendering and a search for sense. *Across Languages and Cultures* 9, pp. 1–15.

Vieira, L. N. 2014. Indices of cognitive effort in machine translation post-editing. *Machine Translation* 28(3–4), pp. 187–216.

Wu, Y. J. and Thierry, G. 2012. Unconscious translation during incidental foreign language processing. *NeuroImage* 59(4), pp. 3468–3473.