23

Translation, the process–product interface and cognition

Silvia Hansen-Schirra and Jean Nitzke

23.1 Introduction: Why do we need a process–product interface?

According to Toury, discourse transfer is inherent in the mental processes involved in translation. From a product perspective, translation depends on a particular manner in which the source text is processed, so that “the more the make-up of a text is taken as a factor in the formulation of its translation, the more the target text can be expected to show traces of interference” (Toury, 1995, 276). Frawley defines translation as “a code in its own right” arising out of the bilateral consideration of the matrix and target codes (Frawley, 1984, p. 168). Baker (1996) pursues this argument and hypothesizes that translation might feature universal properties, which can be found among different translated texts but not in non-translated text. However, the question arises why these translation-specific properties exist. An answer might lie in the translation process itself, since these universal text properties might be triggered by universal processing taking place during translation. 1

From a psycholinguistic perspective, several models of the translation process have been presented (for a summary of earlier models, see Koller, 2004; an overview of more recent models is given in Göpferich, 2008). Most of the models have in common that they split the translation process into a reception and a production phase (see Kautz, 2000; Steiner, 2001; or in machine translation, Vauquois, 1968) or a decoding and a coding phase (see also Schaeffer & Carl, 2013). In addition, some models suggest a further intermediary step, i.e. a transfer step or a conceptualization step (Levelt, 1989). This model has already been empirically tested in Translation Studies (see Carl & Dragsted, 2012; Tirkkonen-Condit, 2005). However, variations thereof have also been presented: Francis and Gallard (2005), for instance, suggest that translators always skip the conceptualization step; Tirkkonen-Condit (2005) assumes that translation is carried out by transcribing the source text into the target language with the help of self-monitoring. Steiner (2001) assumes grammatical deconstruction while simultaneously understanding the source text.

In the following, we argue that the same questions or very similar questions have been independently worked on from two different research branches in Translation Studies: from the product perspective, corpus linguistic techniques have been used to examine translation properties and universals, whereas the process perspective has evolved into the whole area of translation process research. The latter has adopted methods from psycholinguistics and cognitive sciences.
We postulate, however, that we have to bridge the gap between these research branches in order to get a holistic picture of the translation process, one that integrates statistics and explanatory power. Let’s take the dichotomy of normalization and shining through: if we try to relate these translation universals to the process models, we assume that structures are primed without being conceptualized or, in contrast, that a monitoring or inhibition mechanism leads to the mental control of a translation equivalent before it is articulated or written (de Groot, 2011; Levelt, 1999). Accordingly, priming might lead to shining through, whereas monitoring might trigger normalization (Hansen-Schirra, 2017).

In order to investigate these possible correlations, we need to interface product- and process-based research. By doing so, we will get quantitative statistical evidence from the corpus data as well as explanations concerning translators’ strategies, behaviour and cognition. In the following two sections, we discuss the product and process perspectives, outlining relevant methods and their advantages and drawbacks, before introducing the product–process interface in Section 23.4.

**23.2 The product perspective: Corpus linguistics in Translation Studies**

As mentioned earlier, different kinds of translation corpora have been built up in Translation Studies in order to empirically investigate translation properties or universals or to identify translation shifts triggered through specific translation strategies (see Olohan, 2004, or Hansen-Schirra & Teich, 2008 for an overview). The most important methodological considerations will be discussed in the following.

**23.2.1 Corpus designs**

Two types of corpus design are most commonly used in corpus-based Translation Studies: the parallel corpus and the monolingually comparable corpus. Parallel corpora consist of source-language texts and translations of those texts into a target language. They are employed in bilingual computational lexicography, machine translation and translation memories. In translation research, parallel corpora are used to provide information on language-pair-specific translation behaviour, to observe equivalence relations between lexical items or grammatical structures in the source and target languages or texts (see Hansen-Schirra, 2008), and to investigate translation problems and translation mistakes. Some corpus initiatives are moving towards more than one language pair, e.g. the Oslo Multilingual Corpus (OMC) of the SPRIK project (Johansson, 2002).

Monolingually comparable corpora (short: comparable corpora) are collections of translations (from one or more source languages) into one target language and original texts in the target language. Comparable corpora “should cover a similar domain, variety of language and time span, and be of comparable length” (Baker, 1995, p. 23); they have the potential to reveal most about features specific to translated text, i.e., those features that occur exclusively, or with unusually low or high frequency, in translated text as opposed to other types of text production, and that cannot be traced back to the influence of any one particular source text or language.

*Kenny, 1997*

Among the features researchers have posited comparing between translations and texts originally produced in the target language are explicitation, simplification, normalization and conservatism (Baker, 1995, 1996; Hansen, 2003; Laviosa-Braithwaite, 1996; Kenny, 1998; Mauranen, 1997; Teich, 2003). The main application of comparable corpora is the investigation of the
specific (and possibly universal) properties of translations. The fact that translations exhibit linguistics properties that distinguish them from texts that are not translations is sometimes also referred to as *translationese* (see Baroni & Bernardini, 2006).

The most recent type of corpus design is a combination of parallel and comparable corpora. Such a combination will automatically contain a third subcorpus: a multilingually comparable corpus (see Hansen-Schirra et al., 2012). This combination of corpora can be used for cross-linguistic comparison of original texts, cross-linguistic comparison of original and translated texts, cross-linguistic comparison of translated texts, and monolingual comparison of original and translated texts. Methodologically, the primary issue concerning comparable corpora, both monolingual and multilingual, is to define the notion of comparability (see Neumann & Hansen-Schirra, 2013). The favoured solution is to use the concept of register, i.e. functional variation or variation according to situational context. The Translational English Corpus, for example, has followed the design of the British National Corpus in terms of register distinctions.

More recently, empirical translation research has been based on multiple translation corpora. This architecture includes one source text in a given language and target texts translated by several translators into another language. The CRITT TPR DB is such a multiple corpus: several English source texts are translated by multiple translators (professional translators as well as students) into several languages (see Carl, Schaeffer, et al., 2016). Using this corpus, universal translation strategies can be investigated across the language pairs. Moreover, different phases of expertise can be analysed, since novice translation can be compared with expert translation.

### 23.2.2 Corpus annotation and alignment

In Translation Studies, most corpus-based research is carried out on the basis of raw text. Only the encoding of meta-information is a usual practice. More recently, however, the use of annotated corpora is becoming more common, because this is the only way of empirically investigating grammatical and semantic as well as discourse or register features (see Hansen-Schirra et al., 2012). Depending on how abstract the linguistic features to be analysed are, linguistic corpus annotation can be done automatically, semi-automatically or manually. There are a number of automatic annotation techniques. The most reliable ones are part-of-speech (PoS) taggers and morphological analysers, following either a rule-based or a statistical approach (see Hansen-Schirra & Teich, 2008). Also, there are a few treebanking efforts, i.e. annotation of parallel corpora in terms of syntactic structure (e.g. the Sofie Treebank; Samuelsson & Volk, 2005 or the CroCo Corpus; Hansen-Schirra et al., 2012). However, phrase structure and syntactic function as well as semantic annotations have to be corrected or carried out manually.

For the analysis of translation shifts in a parallel corpus, the units of translation (i.e. source-language text units and their translational equivalents) need to be aligned. There are various alignment programs freely available (Hofland, 1996); additionally, aligners are often incorporated in translation memories. The most commonly implemented technique is sentence-by-sentence alignment. Other alignment techniques apply to paragraphs (see Mihailov, 2001) or words (e.g. the word aligner GIZA++; Och & Ney, 2003), the latter being a prerequisite for statistical machine translation. The TreeAligner (Volk et al., 2006) can be used for parallel treebanking, i.e. to align bilingual sentence pairs already annotated in terms of syntactic structure.

Within the context of the CroCo Corpus, a combined query of annotation and alignment enables, for instance, the investigation of so-called “crossing lines” and “empty links” (Čulo et al., 2011): “Empty links” are units in the target text that do not have matches in the source text and vice versa on any alignment layers. The term “crossing lines” is used to denote units whose alignment crosses the alignment of a higher level. Instances of text contained in one sentence or
phrase in the source text but spread over two sentences or phrases, respectively, in the target text can be detected. Such differences probably have implications for the information structure, the distribution of given and new information, or linguistic foci contained in the target text. These structures are very similar to the Cross values discussed in Section 23.4, which clearly shows the need for interfacing product and process research.

23.2.3 Statistics, data analytics and machine learning

Translation corpora can also be statistically exploited, which is the case for machine translation (MT) research. The basic idea of statistical MT is to generate a translation from a parallel training corpus by calculating the most likely equivalent in the target language. Statistical translation models are generated and trained on the corpora with the help of machine learning. Both mono- and bilingual corpora are used to capture the typical linguistic patterns of the languages—the monolingual corpora generate the language model; the bilingual parallel corpora generate the translation model. In addition, statistical MT uses word-aligned n-grams—sequences of words (usually \( n \leq 7 \))—that are assigned probabilities representing how probable the word or sequence is, based on its distribution in the training corpus. Attempts have been made to unite different approaches (usually rule based and statistical) in hybrid systems so that the advantages of the respective approaches can be combined (e.g. Eisele, 2007). The latest approach to MT is the use of neural networks. Neural MT systems try to build one large neural network for translation, while statistical MT systems are composed of many small sub-components. If available, several language pairs are trained together in order to improve the quality and efficiency of the system.

Statistical methods can also be used in order to identify translationese or quantitative patterns in translation corpora (Diwersy et al., 2014; Neumann & Hansen-Schirra, 2013). Baroni and Bernardini (2006) apply text categorization methods and machine learning in order to distinguish translated from non-translated text. Based on support vector machines, they report an accuracy with nearly 90% precision and more than 80% recall for automatically detecting translations. Similar approaches have been proven successful for the automatic classification of translated text (Gaspari & Bernardini, 2008; Koppel & Ordan, 2011; Kurokawa et al., 2009; Nisoi, 2015; van Halteren, 2008; Volansky et al., 2015). Rabinovich et al. (2016) take this kind of research a step further: they use a clustering procedure on the basis of PoS tags in order to classify translated, non-translated and non-native text. Their results show that translations and non-native texts are more similar to each other than they are to original text production. This means that constrained texts, i.e. text varieties that are directly influenced by other languages, seem to share properties, which in turn seems to indicate that universal processes are at play, processes that have an impact on translation as well as non-native text production.

23.2.4 Challenges and problems

When working with multilingual corpora, text segmentation poses several challenges to the comparability of the corpus analyses: For example, French clitics have to be lemmatized in order to render the segmentation between French and German comparable and thus analysable. Above word level, similar problems arise, which concern, for instance, the diverging segmentation of non-finite clauses in German and English. In this case, a possible solution would be to base the recognition of segments on the comparison of the respective functions in each language. In order to investigate the relationship between the source-text corpus and the
Translation: the process–product interface

target-text corpus, the matching text units have to be parallelized. Here, the difficult question is precisely what has to be aligned: whole texts, sentences or smaller units (such as phrases, words or morphemes). For practical or technical reasons, the sentence is usually chosen as the alignment unit, since the automatic detection of sentence boundaries is relatively straightforward. For many research questions (especially the one on the translation unit), this choice is a compromise, the actual focus of the investigation being on the relationship between smaller units. Whereas word alignment is useful for machine translation and multilingual term extraction, this procedure is not sufficient for the analysis of translations, because the translation unit is not reflected in single words. The desirable units for alignment are located between words and sentences and should be flexible enough to cope with typological differences between source and target language. Other alignment problems arise from the fact that different languages have different patterns with respect to sentence length, and that during the process of translating source-language sentences into target-language sentences, the former can be split or merged (see Fabricius–Hansen, 1999).

Another major drawback for corpus-based Translation Studies is the fact that many researchers still rely on untagged, unlemmatized KWIC concordances for their analyses. On this basis, however, only raw text can be found. For instance, irregular language or inflected forms are neglected. A lemmatized or morphologically tagged corpus, where word stems can be explored for the generation of concordances, would offer a higher level of efficiency (i.e. higher precision); a semantically annotated corpus (including synonyms, hyponyms, etc.) would enable more comprehensive querying (i.e. higher recall). Furthermore, string-based approaches are rather lexically oriented. Raw text, concordances or word counts are inadequate for grammatical or semantic investigations, which call for deeper linguistic interpretation. A further problem emerges in connection with cross-linguistic comparisons, as string-based queries fail to take into account typological differences. The analysis of, for example, simplification on the basis of type-token ratio, lexical density and average sentence length, which entails word counts, is not particularly well suited to many language pairs, since typological differences in morphology bias the results (see Hansen & Teich, 1999 for the language pair English–German). Thus, alternative ways of operationalizing the testing of the rather abstract hypotheses both on a text basis and in multilingual environments have to be sought. Syntactic information enables the investigation of explicitation as well as simplification. Thus, deep linguistic annotation (e.g. syntactic, semantic or pragmatic annotation) helps to bridge the gap between the rather abstract hypotheses developed by translation scholars (e.g. on explicitation or simplification in translations) and their realizations in the source and target texts.

For the development of an annotation scheme that meets the requirements of a multilingual analysis, there seem to be two methods that take into consideration the typological characteristics of the involved languages. First, the multilingual corpus is split up into monolingual subcorpora, which are then annotated independently. The second method uses one language as a basis for compiling and analysing a multilingual corpus, whereas the others have to conform. Both methods, however, are rather problematic. The latter forces the adapted language to fit into the system of the language used as a fundament. In the former, questions of cross-linguistic comparison are merely postponed to a later stage and come into play in the interpretation of the data obtained from this research design. In order to avoid both problems, the annotation scheme has to abstract from the level of language-specific realizations to functional categories that are comparable across languages. In order to keep inter-annotator agreement and thus the annotation quality to an acceptable level, annotation schemes and alignment rules have to be very precise and transparent (see Brants & Hansen, 2002).
23.3 The process perspective: Translation process research

From the process perspective, analyses of the source text and the translation product are also very important. They are necessary to outline what might be happening in the translator’s mind. To analyse cognitive translation processes, however, the following methods are well established in translation process research (see Göpferich, 2008; Krings, 2005).

23.3.1 Report data

Think-aloud protocols (TAPs) are used to record translators’ thoughts during the translation process. Translators can be asked to verbalize their thoughts directly during the translation or retrospectively; the first variant, however, is more common. The transcriptions of these verbalizations are called think-aloud protocols. Using TAPs has various advantages and disadvantages. Among others, one major disadvantage of immediate verbalization is that studies have shown that verbalization changes the thought process (see Jakobsen, 2003) and therefore may also change the translation process. TAPs that are produced retrospectively do not change the translation process, because the translator is asked about specific translation units only after the whole translation has been produced. The problem is, however, that a lot of valuable thoughts might get lost between translation production and verbalization. Here, it is helpful to use a screen recording of the session—maybe with eye-tracking data—to help the participant recall the passages of interest. Further, only thoughts that are conscious can be uttered, and a high cognitive load during the task might prevent participants from verbalizing their thoughts. However, it is one of the few methods that actually purport to reproduce what is going on inside the heads of the participants, even if not completely (Jääskeläinen, 2010).

Questionnaires are usually distributed in a written form—either on paper or electronically—and contain a set of questions. These questions can be open—the participant decides which questions to answer and in how much detail—or closed—the participant can choose from a set of answers. Mixed questions elicit a number of possible answers. Participants can, however, add their own answers. Closed questions are much easier to assess, which saves time and money. Open questions, however, deliver more extensive and more differentiated data (Klöckner & Friedrichs, 2014). Compared with interviews, questionnaires are easier to distribute, and participants might be more willing to fill out a questionnaire whenever they have time than to schedule a date with the interviewer. Further, questionnaires are more discreet and more anonymous. On the other hand, participants must be able to read and write, which excludes some potential target groups (but this is not usually a problem in translation process studies). Written answers will probably be shorter and less detailed than answers in an interview. Finally, a questionnaire does not allow personal contact, so that participants are not able to ask questions and the examiner cannot get an impression of the person (see Döring & Bortz, 2014, pp. 398–399).

The requirements for a high-quality questionnaire are that the participant is able to answer the questions and that it is objective, valid and reliable. However, even a very sophisticated questionnaire can cause non-responses and different response qualities due to the individual characteristics of the participants. Furthermore, participants tend to follow certain response strategies, e.g. they prefer extreme categories or medium categories, they prefer the first or the last response choice (also called primacy or recency effect), or they answer according to social conventions (see Reinecke, 2014, pp. 612–613).
23.3.2 Behavioural data

Keylogging software allows the researcher to analyse the text production process and the associated mental processes. All key (and mouse) activities are recorded during text production, including typing processes, special key combinations and deleting activities. Further, pauses are recorded, which can elucidate reading processes and the segmentation of the text, which is done subconsciously by the participant and might indicate text passages that require high cognitive effort (see Alves & Vale, 2011). However, one has to keep in mind that the data only allow us to speculate on mental processes of participants engaged in text production. Other methods, especially eye tracking, can help to interpret participants’ behaviour, e.g. during pauses (see Jakobsen, 2011, 37–38). Recording keylogging data with a keylogger is unobtrusive, because the program runs in the background and, hence, the recording process is not noticeable (see Carl, 2012, p. 4108).

The term eye tracking refers to the methodology with which human eye movement (called saccades) and fixations can be recorded and quantified. The basic assumption is that the human pays attention to the point (s)he fixates on (a concept also known as the eye-mind assumption, introduced in a reading study by Just & Carpenter, 1980), although this is not always the case. The eye movement data give us promising and important hints about what is going on in the translator’s mind, although we can only interpret the data and cannot be entirely sure what is happening in the black box. This research method, however, brings not only advantages but also challenges to translation process research. For instance, the kind of eye-tracking system has to be considered. A remote eye tracker is considered ecologically more valid, because the participant can move comparably freely in front of the computer screen, while head-mounted eye trackers and eye trackers with chin or head rests produce more accurate data. Eye-tracking glasses, on the other hand, liberate the participants from the screen. In contrast to questionnaires, think-aloud protocols and keylogging technology, which involve only very low costs if the experimenter has a PC or laptop and an Internet connection, eye tracking requires expensive hardware and software. Further, eye-tracking experiments must be conducted in a consistent environment, e.g. with similar lighting conditions for all participants (see O’Brien, 2009, pp. 251–254; on detailed information on requirements for a suitable eye-tracking lab, see Rösner, 2016). The texts should be no longer than about 300 words in eye-tracking studies so that scrolling does not become necessary, because this makes the data harder to assess. Similarly, the texts should be in a font size of 16 or 18 and displayed with at least 1.5 line spacing so that the eye-tracking data can be mapped correctly (see O’Brien, 2009, pp. 261–262).

23.3.3 Neuroscientific data

Translation Studies have only in recent years slowly begun to use neuroscientific methods to ultimately tackle the problem of what is going on in the translator’s “black box” while translating. These methods are typically used in very controlled experiments and cannot (yet) be used in authentic translation tasks. Electroencephalography (EEG) is used to record electrical activity in the brain. Depending on what needs to be measured and how precise the recordings need to be, 16 to 256 electrodes are distributed on the head. Artefacts in the EEG signals can be produced by external influences like head movement, blinking or other muscular activities. These artefacts cannot easily be avoided and have to be eliminated from recordings, if possible, before the signals are analysed and interpreted (see Freeman & Quiroga, 2013, 5–6). The electrical activity in the brain produces oscillations, which have been associated with certain functions and pathological conditions of the brain. However, not only oscillations but also
typical patterns can be studied with the help of EEG signals. An example of such an event-related potential (ERP) would be the N400. This amplitude peaks negatively about 400 ms after a participant has been presented with a stimulus. The N400 is widely acknowledged as a measure for semantic processing. If, for example, a sentence is presented that includes a semantically nonsensical stimulus, the negative amplitude will be easily visible 400 ms later (see Kutas & Federmeier, 2011). Translation Studies has used EEG, for example, to investigate priming, monitoring and inhibition (see Oster, 2018), cognitive effort in backward translation (García et al., 2015), conceptualization (e.g. Grabner et al., 2007) and expectancy violations (e.g. Elmer et al., 2010)—for an overview, see Hansen-Schirra (2017).

The applications of functional magnetic resonance imaging (fMRI) include all areas of brain imaging and have become a very important tool for neuroscience research, both clinical and cognitive (see Faro & Mohamed, 2006, pp. v–vi). Kim and Bandettini (2006, p. 3) observed “that functional brain mapping is possible by using the venous blood oxygen level-dependent (BOLD) magnetic resonance imaging (MRI) contrast”. The BOLD method depends on the level of deoxyhaemoglobin, which can be seen in the signal intensity of magnetic resonance images when the level changes and can hence be used for human brain imaging. Compared with other brain imaging methods, fMRI is especially useful to show language areas in the brain, because it is non-invasive and produces images of good quality and with good localization (among other benefits). It is rather difficult to assign brain regions to single language processes such as phonetic, semantic or syntactic processes, because they often work together. Carefully selected research designs with contrasting conditions can, however, help in tackling these problems (see Binder, 2006, 245–248). Ahrens et al. (2010), Franceschini et al. (2003) and Kalderon (2017) were the first researchers to use fMRI to localize and visualize activated areas during translation and interpreting.

### 23.3.4 Challenges and problems

Despite the differences discussed with respect to the individual methods mentioned, they all have in common that the participants for a study must be chosen carefully. Although professional translators are often considered more valuable, as they have practical experience, they are harder to convince to participate in a study and often expect financial compensation, as they miss (part of) their workday. Students, on the other hand, are easier to recruit, as eye-tracking studies in translation process research are usually conducted at a university, and the study might even be credited in classes. However, students exhibit different cognitive processes and behavioural patterns than professionals, which might in turn affect the generalizability of the results. Further, some studies, like questionnaire studies, may be carried out at home, while other studies, especially those that require special equipment, as most eye-tracking, EEG or fMRI studies do, can only be conducted in special laboratories. In addition, one question is whether the participants will commit to an equal degree when they invest their free time into participating in studies as opposed to when they are paid or rewarded in a different manner for the task. Additionally, limited funding might restrict the number of participants that can be recruited for the study. However, it is doubtful whether small participant numbers, e.g. 12 participants or even fewer, can return generalizable and well-balanced results. Nonetheless, these studies are valuable to build hypotheses for larger studies. Finally, the professionalism of the participants has to be addressed, as not every translator with a degree in translation is equally capable of all tasks. Some issues may also occur which disqualify the participant for the study, e.g. typing skills, language competence or ability to follow instructions, or if the participant feels intimidated, judged or pressured during the session (see O’Brien, 2009, pp. 254–259, p. 262).
Another problematic issue concerning TPR studies concerns the authenticity of the translation situation. While some methods are quite authentic in their experimental setup—i.e. the participant can use a computer or laptop and online research for the translation of authentic texts—others are not natural at all, e.g. those involving an EEG cap or an fMRI scanner. The latter usually focus on single-word translation or translations, which are not written or verbalized at all, while the former also take place during a phase of observation. These experimental conditions are needed in order to figure out which effect is triggered by which stimulus. Furthermore, not only are many participants needed, but so are many repetitions of the stimuli, to improve statistical validity and significance. Thus, the more controlled an experiment is, the easier it will be to see a clear stimulus–effect relation. So, there is always a trade–off between authenticity of the translation situation and effect size concerning a given hypothesis.

23.4 The challenge of interfacing the two perspectives

Data triangulation—i.e. linking two or more sources of data, researchers, methodological approaches, theoretical ideas or analytical designs so that the different advantages of each methods can be exploited (see Thurmond, 2001, pp. 253–257)—has become more and more important in Translation Studies (e.g. Alves, 2003). Triangulation has the advantages that the research data become more reliable, inventive approaches are developed to comprehend and interpret research hypotheses, existing theories might be challenged or confirmed, and a phenomenon can be better understood. Every triangulation approach on its own, whether it is data, researcher, methodological, theory or analysis triangulation, has individual advantages, but also disadvantages. In general, these include

- an increased amount of time needed in comparison to single strategies,
- [...] difficulties of dealing with the vast amount of data,
- [...] potential disharmony based on investigator biases,
- [...] conflicts because of theoretical frameworks, and
- [...] lack of understanding about why triangulation strategies were used.

Thurmond, 2001, p. 256

In short, triangulation is valuable to gather findings from different perspectives that are complete and confirm each other, and to strengthen the findings. The researchers, however, must be able to explain why they used the triangulation method and why it was necessary. Many studies have adopted this approach in translation process research in recent years, especially concerning data triangulation. The danger, however, is that, among other possible problems, the amount of data becomes overwhelming. Some solutions on how to deal with large data sets might be to work in research teams, so that either the same research topic is analysed together or different researchers analyse various hypotheses on the same data (O’Brien, 2009, pp. 260–261, p. 264).

What we suggest here exceeds the triangulation of different perspectives, since we propose to interface different research branches, namely that of corpus linguistics with that of TPR. There are, however, two ways of achieving such interfacing:

- different data sets are used with different methods to answer one research question, or
- the same data set is used with different methods to answer one research question.

In the following, we will present both alternatives on the basis of selected examples.
23.4.1 Interfacing different data sets with different methods

The first studies in which different data sets were interfaced using different methods can be found in Hansen (2003) and Alves et al. (2010). The former examined the relation between normalization and translation expertise; the latter the relation between translation units and grammatical shifts occurring during translation. However, in both studies, the process-oriented task was rather example based and could not have been tested with statistical methods. The generalizability of the results is therefore limited. However, these studies can be regarded as pioneer work, given the methodological principle they adopt of interfacing product and process research.

For a more empirical approach, let’s get back to our dichotomy of normalization vs. shining through mentioned in Section 23.1: Tirkkonen-Condit (2005) argues that literal translation, which may result in shining-through effects, is a default translation procedure, which is cognitively preferred to others. Chesterman (2011) and Halverson (2015) reintroduce the concept of literal translation, assuming that entrenchment effects strengthen the co-activation or priming effect of linguistic patterns and thus reduce the cognitive load during translation for literal renderings (see Schaeffer & Carl, 2014 for an empirical operationalization). Tirkkonen-Condit (2005) assumes that translators constantly monitor production, and as soon as a problem is encountered in their default translation routes, they stop the literal translation process and try to find a better solution. The monitor model has been empirically tested by Carl & Dragsted (2012).

The process-oriented continuum between monitoring and priming, i.e. literal translation, could be another way to perceive the product-based dichotomy between normalization and shining through. The monitor model, however, still exhibits some shortcomings. It is, for example, not precise enough to determine which factors influence priming and which linguistic levels are affected over others. In order to further clarify these open questions, Hansen-Schirra et al. (2017) investigated cognates (translation equivalents which share a similar form). Cognates are relatively easy to control in experimental settings and can be investigated in many language pairs. Using EEG methodology, Christoffels et al. (2006), de Groot (2011) and Oster (2018) show that cognates seem to be pre-activated during translation. This cognate facilitation effect indicates that priming takes place while participants are decoding the source text. On the other hand, Kußmaul (1989) argues that monitoring takes place while translators are translating cognates. In order to find out which factors trigger priming or monitoring processes, Hansen-Schirra et al. (2017) exploited large multilingual corpora to quantify cognate usage patterns, on the one hand, and carried out controlled experiments to explain translation-inherent factors for cognate translation, on the other. Using the Google n-gram viewer, they trace the impact of societal and technological development on the usage of selected cognates. They show cases where a cognate was first used in one language, but due to common language roots, it was accepted in other languages as well. In addition, typical preferences in terms of cognate usage could be identified for different text types in the German DWDS corpus. It seems to be the case that some text types, such as academic texts, are more receptive to cognates than others (e.g. newspaper texts). When comparing translation with post-editing corpora, they found that translations from scratch and post-edited target texts show similar cognate and non-cognate usage; however, the variety in non-cognate lexical choices is statistically higher in translations from scratch than in post-edited texts, for which the machine translation output was relied on in most cases. This shows that the usage of computer-aided translation may also have an influence.

Complementing the corpus results with experimental methods, Hansen-Schirra et al. (2017) contrasted single-word translation of cognates with translating them with context. The analysis showed that the use of cognates in translations is dependent on the context of the translation.
In general, the participants chose a cognate less frequently when they were translating a whole text than when they only had to find German equivalents in a word list. This might indicate that the cognate translation is the most obvious and the easiest without context, because the cognate is similar not only in meaning but also in form. When a cognate is embedded in context, however, the surrounding co-text of the word triggers different lexical choices. Another translation experiment included translation students with different levels of expertise. The comparison here suggests that monitoring or other control mechanisms develop with increasing translation competence, since the number of cognates decreases for more experienced translators. In addition, other studies have shown that the number of cognates in translations varies significantly depending on other factors, such as status of the respective languages (Vintar & Hansen-Schirra, 2005), translation mode (Oster, 2017) or exposure to media in the mother tongue (Oster, 2018).

In conclusion, it can be argued that some of the factors that might influence the production of cognates can only be isolated with corpus-based research, while others have to be approached from a more process-oriented perspective. A combination of both comprehensively shows how the translation of cognates can be predicted against the background of different constraints.

23.4.2 Interfacing the same data set with different methods

When interfacing the same data with information from several data collection methods, corpus enrichment becomes more complicated, since user-activity data is added to the text corpus. This means that each word, phrase or sentence is annotated not only with linguistic information, such as PoS tags, phrase structure, co-reference resolution, etc., but also with data concerning the translator’s behaviour during the translation process. More concretely, data on eye movements, i.e. fixation times, saccades, regressions, etc., have to be mapped onto the source and target units. Information from the keylogs, i.e. on the production processes of words as well as pauses, also has to be added. The representation of data including all kinds of annotations and alignments can be quite complex. If the baseline is the word (it could also be the syllable or a time-stamp for spoken data), there are several alignment and annotation relations:

- product annotation with linguistic information (PoS, phrase structure, semantic relations, morphology, etc.)
- product alignment with source text (word alignment, sentence alignment, etc.)
- process annotation with production units from keylogs (production history, pauses, etc.)
- process alignment with eye-tracking data (fixation duration, fixation count, etc.)

While product annotation and alignment are statistical in nature, user-activity data from the process constitute dynamic information. Some of the linguistic annotations and alignments may be hierarchical in their structure (e.g. syntactic dependency trees), while others reflect temporal progress (e.g. eye data). Both phenomena complicate the corpus representation and the holistic integration of all data layers as well as querying procedures and corpus analyses. Moreover, the data may be further processed, resulting in more specified metrics, e.g. the Cross value (explained later), which in turn may again be added as annotation to each word of the corpus. This makes the corpus itself a dynamic resource.

The first dynamic corpora including user-activity data deal with individual language pairs and include either keylogging (e.g. Alves & Gonçalves, 2013; Jakobsen, 2002; Serbina et al., 2015) or eye-tracking data (e.g. Jakobsen & Jensen, 2008). The CRITT-TPR database is the first dynamic corpus for several language pairs and includes eye-tracking and keylogging data in addition to linguistic annotation. This multiple translation corpus includes different translation tasks (from
Hansen-Schirra and Nitzke

scratch, post-editing MT, etc.) by professional and student translators (see Hvelplund, 2011). While
some data categories represent “pure” eye-tracking and keylogging data, other parameters are
provided which present processed data (e.g. the Cross values explained later). Finally, there is
also additional information on the source and target-text units, such as PoS tags (see Carl &
Schaeffer, 2013 or Carl, Schaeffer, et al., 2016; these papers also provide a detailed description
of the parameters contained in the database). Schaeffer & Carl (2014) introduce a metric that
operationalizes the above-mentioned literal translation hypothesis. Further, they evaluate the effort
that is necessary for non-literal translation. To do so, they use the gaze behaviour of translators
for different language pairs from the CRITT-TPR DB. Different lexical realizations of source
words in the target language are counted, and for some words, higher variation is detected than
for others. On a syntactic basis, crossing word alignments (similar to the crossing lines in Section
23.2) are introduced: the metric computes the Cross values for single words, based on their pos-
ition in the source and the target language. Depending on the point of view, the Cross value can
be realized from the source text as a reference and the target text as output (“CrossS”) or the other
way around (“CrossT”). The smaller the Cross value, the more similar the texts are in terms of
structure; when the Cross value is high, syntax varies significantly. The authors show that higher
Cross values strongly correlate with total reading time on source and target words and, therefore,
prove that high syntactic variation takes more effort to produce. Furthermore, a strong correl-
ation can be found between production time of the target-text word and number of alternative
translations. “With few choices post-editors are quicker than translators, but this distance decreases
as the number of translation choices increase” (ibid., p. 34). Additionally, a strong correlation was
detected between total reading time on target-text word and number of alternative translations.
This strongly supports the literal translation hypothesis by attesting higher cognitive load for non-
literal translations—taking lexical as well as syntactic entropy into account. Bangalore et al. (2016)
investigate the role of co-activation of languages in translation and its influence on the translator’s
behaviour, combining syntactic annotation with user-activity data.

23.5 Concluding remarks and outlook

In this chapter, we have:

• clarified the need for interfacing process and product data. However, we have also explained
  why it is not sufficient to triangulate different methods; rather, we need to combine different
  methods from different research branches in Translation Studies, i.e. corpus linguistics and
  translation process research.
• introduced the most common methods within the two research branches and discussed their
  strengths but also their shortcomings and problems.
• given examples of successful product–process interfaces—with a common data set or without.

We have shown that corpus linguistics leads us to the quantification of translation patterns.
However, in order to explain the quantitative phenomena, it is necessary to complement them
with cognitive experiments. For instance, product analyses may identify normalization or
shining through, but only on the basis of process-based research can we explain how and why
translators decided upon the strategies that trigger exactly these phenomena. The combination
of corpus-based work on translated texts (i.e. product-oriented research) with experimental
studies of translators’ behaviour (i.e. process-oriented research) seems to be promising, since it
enables not only the quantification and interpretation of translation properties and strategies
but also their explanation concerning cognitive processes and resources. Recent advances in
corpus architecture and corpus querying (including multi-layer annotation and alignment) and increasing incorporation of methods from psycholinguistics and cognitive science into process-oriented research (especially eye tracking and keystroke logging) point to a desired combination of corpus studies with a more direct, experimental insight into processing efforts.

Future work will include the integration of speech data such as translation dictation or sight translation, on the one hand (see Carl et al., 2016b), and deeper analysis as well as the meta-annotation of user-activity databases, on the other. With multimodal annotation tools, such as the web application developed by Alves and Vale (2009), it is possible to manually annotate the user-activity data and to query these corpora. Such tools facilitate translation-oriented annotations, such as the explicit labelling of translation procedures or the annotation of explicitation. On this basis, correlation between these annotations and the eye-tracking or keylogging data can be calculated. It can be clarified whether, for instance, explicitating ambiguous structures will influence source-text reading times.

Another development will be the integration of EEG and fMRI data into user-activity corpora. So far, these methods have been adopted from cognitive sciences in order to identify biological correlates of cognitive and motoric processes. However, they have not yet been integrated into multi-method approaches to translation. Focusing on single words as stimuli, or not being able to actually produce the translation orally or in written form, poses several problems for the investigation of cognitive translation processes: As translators usually translate whole texts in context and with respect to a given *skopos*, it is difficult to study complex translation processes or strategies using EEG or fMRI technology as described earlier. Decoding the source text and encoding the target text involve such a wide range of stimuli that it is difficult to isolate a single stimulus during an authentic translation task. This means that studies in which single-word translation is used or in which the translation has to be produced mentally are not ecologically valid, because they neglect the complex problem-solving mechanism employed during translation. Although the context problem has been relativized by van Hell and de Groot (2008), who found similar effects for context-free vs. sentence-context conditions, there is still a huge gap between existing translation-oriented theories and models and their operationalization with methods from cognitive science (see Göpferich, 2008). However, Nagels et al. (2013) integrate authentic text reception and production into fMRI methodology, which seems to be very promising for translation process research. Finally, it can be argued that it is beneficial to triangulate data collected in manipulated studies, on the basis of which clear effects can be derived with sentence-context studies that corroborate the results gained from the controlled setting. By interfacing different methods with different data sets, experimental control can be complemented with ecological validity.

**Note**

1 A critical discussion of the existence of translation universals can be found in Mauranen & Kujamäki (2004), Malmkjær (2005) or House (2008). Hansen-Schirra et al. (2012) reframed the concept and more cautiously call them translation-specific properties.

**Further reading**


This collection of articles focuses on translation process research but also includes the integration of translation product analyses. The data triangulation is reflected from a theoretical and a methodological point of view.

This volume focuses on the investigation of keylogging and eye-tracking data. Using multi-method approaches, “old” questions concerning, for instance, priming can be re-analysed.


This book focuses on the data-driven modelling of translation and interpreting processes. By interfacing process and product data, more comprehensive models and theories can be formulated.

References


Translation: the process–product interface


Hansen-Schirra and Nitzke


