6. Conclusions

Upon investigating the previous studies pertaining to the process of translation as discussed in the chapter on the Review of Literature, scholars such as Wolfgang Loscher, Susanne Gopferich, have come together and attempted to carry out empirical investigations into translation both as a process and as a product from within the field of Translation Studies. This has resulted in the formulation of research methods such as audio recording of Think Aloud Protocols, keystroke logging, eye tracking to capture the gaze points, etc. that involve technology and other qualitative methods like retrospective interviews and questionnaires. These scholars have borrowed these methods from various other fields and have tailored them to suit the needs of Translation Process Research. They have designed their own experiments for collecting their own data by adopting a multiple method approach to elicit data. They have preferred not to use the conventional parallel corpora, which could have been due to the absence of suitable corpora to study the process of translation as the findings of the earlier chapters indicate that the data sought by the scholars working on process research is completely different.

Although the results obtained through such experiments and methods are accessible to the research community, the modes and means adopted to arrive at these findings are not available, and this turns out to be a handicap for them to validate or further build on the previous studies. In this context, many researchers such as Mona Baker (1993), Maria Tymoczko (2010), Fredrico Zennettin (2011) etc. have advocated the necessity to make these modes and means accessible to the research community by adopting corpus methods that have contributed in a substantial way to contrastive and comparative studies in the fields such as linguistics.
The corpus based research projects have challenged and proven that the theories and generalizations made by the earlier researchers were intuitive through validations based on the corpus data available as demonstrated by Biber (2010:167). In addition to being the means to validate, corpus also has made it possible for the researchers to come forth with completely new theories through corpus driven studies. However, it may be said that the potential of corpus methods has not been exploited to their fullest in the context of researches in Descriptive Translation Studies. The corpus methods are still in their evolutionary stages as there could be many more possibilities that are yet to be uncovered by innovative and creative ideas of future researchers.

Some corpora, e.g. English Swedish Parallel Corpus (ESPC) manual on their website, claim Translation Process Research as one of the objectives. However, many process researchers chose to create their own data, as the data sought by them differed in terms of the nature of information contained in existing parallel corpora. While the parallel corpus focuses on the equivalence of words and other linguistic elements like metaphors etc., the data created by the process researchers included information relating to the problems faced, the solutions provided, resources used and strategies adopted by the translators. The data has been used exclusively to study the patterns pertaining to various aspects of translation process. Further, most of process researchers have employed multiple methods for eliciting data in different combinations.

This may have been done for two reasons:

1. Translation as an activity is a complex mental process involving various intrinsic linguistic and cultural processes that occur in the translator’s mind that are not directly accessible to researchers through visual observations.
2. The results may be correlated and validated through “triangulation” of findings suggested by data collected from different sources and methods in order to avoid erroneous and intuitive generalizations.

Most commonly used data elicitation methods, in addition to the source text and target text are:

- Think Aloud Protocols
- Audio recordings
- Video recordings
- Keystroke logging using keystroke logging software such as Tanslog
- Activity logging software such as PROXY
- Retrospective interviews while replaying video, audio and keystroke recordings
- Questionnaires

This may imply that the existing parallel corpora may fulfil few research goals with regard to Process Oriented Study of translation as the data contained is only in the form of source text and target text and is only retrospective in nature. On the other hand, the process researchers prefer data elicited in real time by means of multiple methods that help them to get a clearer insight and at the same time cross verify their findings through triangulation, in addition to the comparative analysis of the source and target texts as a basis for theorization. Therefore, it may not be unreasonable to say that a corpus that is intended for process studies needs to integrate data and information elicited using multiple methods, in other words, adopting a triangulative approach. Doing so may benefit the researchers in following ways:

1. Saving effort, time and money involved in eliciting separate data for each research. For example, scholars interested in studying a particular
phenomenon pertaining to process of translation between Telugu and Japanese pair can use this kind of corpus, without needing to create their own data.

2. The element of subjectivity can be avoided as the data available is not created for investigating a specific research question (objective neutral).

3. Reliability of the findings may increase when the other scholars are able to replicate and quantitatively validate the findings.

4. Since the data also includes the source text and target text, it can be used as a conventional corpus of parallel aligned texts as well.

5. Integrating this new Corpus approach into Translation Studies may further catalyze evolution of newer methods in Process Oriented Studies.

6. Indirect benefits may include convergence of research objectives as researchers from various other fields of study such as Computational Linguistics, Corpus Linguistics, and Translation Studies etc. come together.

Looking at the quantitative and qualitative Process Researches in Translation Studies as seen in the chapter on Review of Literature; and considering the benefits and insights that incorporation of this approach can provide with regard to the qualitative and quantitative researches as discussed in the chapter on triangulation and software evaluation, Triangulative approach seems to play a vital and integral role in process researches in Translation Studies. Process researches till now have captured the data by combining different methods. These methods can broadly be categorized into those that use technology and those that do not involve technology.

Keeping the methods that are dependent on technology in mind, three types of software were evaluated. All three of them differ from each other in terms of kind of features that each software possesses. After evaluating the relative advantages
After looking at the needs of process researchers that were assessed after a review of previous researches and the solutions that corpus studies can provide, a model corpus was suggested. The suggested model differs from the conventional parallel corpora in following ways:

- The suggested model contains versions of translated target text. For example, in Page 1 the translator has mistranslated the word “mama” as “so so” however, towards the later versions he has corrected it to mother which is the right translation. This would not have been captured in the conventional parallel corpus as it only contains the final version of translated text
- The source text, versions of target texts and the transcripts of the verbal reports are aligned at sentence level.
- Audio recording of Think Aloud Protocols, Video recordings of the subject and the screen recording are all available in a single file.
- These recordings are available separately for each version
- Hypothetically, the relevant parts of the texts can be retrieved using descriptions of problem etc. through the tags of the transcripts

Limitations of the Model
Every model comes with its own shortcomings. Some of the limitations of this model that can be thought of at this point of time are as follows:

- Memory space is an aspect directly related to corpus as it needs to store a lot of data. The proposed model needs to store video and audio recordings in addition to the conventional data; therefore the file sizes will be large. Further, the storage of translated versions is likely to increase the size of the data to be stored. However, as seen in the past, technology has
continuously evolved providing faster processing and larger memory space; it is quite likely that these problems will also be overcome in times to come. One immediate solution possible is to store the video in lower resolution in formats such as .wmv or .avi, as is done in this research.

- The current search tools are directed towards retrieving information according to search categories that have been created by scholars of traditional linguistics or other fields that make use of corpus. However, there may be very few or no search categories that the researchers of Translation Process may need for retrieving required data.

- Compilation of conventional corpora involves collecting the source and target texts that are readily available and does not involve any special efforts to create the data. However, the data for this kind of a corpus will not be readily available and have to be created afresh which might need a conscious effort by the research community.

- Availability of subjects might prove to be an issue in view of the following reasons:
  1. Willingness of the subject
  2. Availability of subjects especially in cases of the language pair that are rare. In case of this study there was only one subject available

- Unknown time and cost factors involved in building the said corpus.

**Further Scope for Research**

- The requirements of researchers in Translation Studies may be studied thoroughly from the point of view of extending corpus methods in order to understand the kind of information the translation process researchers
require, the methods they adopt to elicit and analyze data. This may be useful for creating standardized corpus design and methods intended for Process Research.

- New tagging categories can be created after thoroughly studying the strategies adopted by the translator and the phenomenon observed while performing a task. These can be later used to tag the transcripts and enable retrieval of data using these categories as search criteria.

- The existing corpora stores information in text format that is tagged using XML, since the corpus model suggested as part of this research also stores the data in text format, it may be assumed that conventional methods can be applied. However, there may be a need to modify the existing tools or create new tools for aligning texts and searching corpus suiting the process research.