

Book Review

Research Methodologies in Translation Studies

Mehrnoosh Pirhayati *

M.A. in English Language Translation, Department of English Language, Science and Research Branch, Islamic Azad University, Tehran, Iran.

This book review is very important and necessary for those who engage in Translation Studies (TS), specially translation scholars who want to become more familiar with translation methods and have a good performance during their scholarly investigation and research. This book was written by Gabriela Saldanha and Sharon O'Brien (2013) and published by Routledge. It is divided into seven chapters and started from introducing research models and the content of the book and ended by chapter seven entitled as "Conclusion: The Research Report". All chapters ended by a summary that the authors briefly re-described and re-explained the aims and purposes of each chapter. In fact, this book can be viewed as a dictionary of the research terms of TS and also as a guide that shows the strengths and weaknesses of different scientific methods of research in this field. So, the purpose of this book review is to mention the important key terms and also the strengths and weaknesses of the used methods in this scientific field.

As previously was noted, chapter one was referred to introducing the scope and purpose of this book. In chapter two some key terms as the basic and elementary knowledge that are necessary for scholars were brought and described by the authors. They described the key term of ontology as the way that social world can be seen and relates to the reality of the social phenomena that constitutes the social world and gives us three ways of seeing the world that are: objectivism, realism, and constructivism. Epistemology was another term that was described as the theory that lets us to know things on the basis of its rules and principles that generally three approaches can be attributed to it including, positivism, interpretivism, and realism. The realism gets the location that is as between the objectivism and constructivism views. The other elements as the key term was framework that was explained as a set of rules and principles that scholar or researcher by employing it can be able to gather knowledge and scientifically or systematically view its particular domain.

A concept is an idea while theory is a set of ideas and concepts that can be used as the means of understanding. Method is research technique, while a methodology is defined as the science of method (Sealy, 2010, pp. 61 as cited by Saldanha & O'Brien, 2013, pp. 13). After, two general types of research were explained as those researches that can be

subcategorized as inductive researches, and those can be referred to as deductive researches. Induction moves from little and small scientific details to developing or producing a general theory or hypothesis, while deduction is for testing a theory or hypothesis through knowledge and data. Other types of research are: empirical research, explorative research, conceptual research, applied research, evaluative research, action research, and ethnographic research. Hypothesis can be casual, which indicates to the fact that change of the item of X can lead to the change of the item of Y, or associative that proves the effect of one concept on another one. It should be noted that, hypothesis is expression and it is not question or phrase and only asserts the relationship(s) between two or more concepts (Saldanha & O'Brien, 2013, pp. 18).

A hypothesis can be descriptive, explanatory, or predictive as empirical hypothesis, or interpretive as conceptual hypothesis. After, the key term of data was clarified by the authors that can be the two types of primary data and secondary data. Primary data are collected by the researcher and the secondary data are those that are collected by the other researchers. Structured data or quantitative data are those that are statistically analyzed and are numerically represented while unstructured or semi-structured data relate to the qualitative approach and cannot be categorized as structured data. Quantitative data were categorized as categorical scale data or nominal data, ordinal scale data, interval scale data, and ratio scale data. The difference between ratio data with interval scale data is ratio data, unlike the interval scale data have an absolute zero point. Pilot study is defined as a small-scale study that gives to the researcher the knowledge of the usefulness of the selected methods of analysis and also the extent or degree of gathering data that can establish the credibility of research, but only with regards to this point, the size of data.

Three approaches were enumerated by the authors as the quantitative approaches that relates to positivist epistemological position, the qualitative approaches that relates to the interpretivist position, and mixed-methods approach was described as when the researcher uses both quantitative and qualitative approaches or methods to analyze or collect data. After, the unit of data was foregrounded and as these two different types, macro-level data and micro-level data was introduced. Macro-level data are the result(s) of analyses that are collected as data from organizations, systems, countries, while micro-level data are those data that are seen at the level of individual, word, or text. The next was the definition of the term, variable that was defined as something that varies and can be measured, and it can be independent and dependent. Independent variable can be manipulated to see the influence(s) on the dependent variable, while independent variable is the core concept and the researcher tries to assess it and can be changed when it is exposed to different or varying treatment. The other categories of variables are profile variables that relate to "the form of a translation" and context variables that relate to "the aspects of the translation context and its consequences" (Saldanha & O'Brien, 2013, pp. 27). Next, the term of validity was mentioned as when the results of the research, as closely as possible, match the real state of the world, or a research that is produced based on regarding the scientific and moral aspects as the two parameters of defining validity.

The other definition of validity was that of Frey et al (1999, as cited by Saldanha & O'Brien, 2013) that based on their approach three categories as threats were related to validity including, internal threats, external threats, and the threats that consider measurement. Internal threats relate to the researcher, research participants or how the research is conducted, and external threats relate to replication, and ecological validity, and threats relate to measurement can be due to the selected method for sampling and the size of samples. In their book different types of sampling were mentioned and explained as random sampling or probability sampling or the gold standard of sampling, systematic sampling, stratified sampling that may random or systematic, cluster sampling, stage sampling that occurs after taking a random sample from a large population. Then, by applying this kind of sampling a smaller sample from a large group can be studied, purposive sampling that the selection is based on the pre-defined critical parameters and commonly used in the corpus-based studies, convenience sampling that relates to drawing sampling from that part of population that is close to hand and snowball sampling or chain sampling or referral-chain sampling, or referral sampling is a nonprobability method of sampling that sample grows like a snowball.

After, they defined reliability, as reproducibility or replicability of a particular research by the other researchers and reaching the same conclusion under the same conditions. Generalizability was another key term which was defined by them as the potentiality of research that based on it the result of the research can be attributed to the large number of a particular population or the researcher can be able to make claim about a large number of population based on regarding this potentiality of its research. Although a research may not be generalizable under regarding these three conditions of 1- for generating a hypothesis or 2- for testing the viability of a framework, and 3- in exploring the questions that were built up by how and why, but they are considered as having the quality of research. They stated that the subjects of validity, reliability, and generalizability sometime were played own by the researchers in their qualitative research and they foregrounded four concepts that were defined by Lather (1986, as cited by Saldanha & O'Brien, 2013) including, face validity as covering the subject that was supposed to be investigated or inquired in research and participants are for example agree that the method or the prepared test looks like it is going to measure what is supposed to be measured in that particular research, catalytic validity that increases self-understanding and self-direction of the researcher, since shows the reality-altering impact of the research process, triangulation of methods, data sources, and theories, and reflexive subjectivity.

On the other hand, they foregrounded the approach of Wood and Kroger (2000, as cited by Saldanha & O'Brien, 2013), the concept of warrantability as the extent of trustworthy of a research as the determinant of evaluating the quality of the qualitative research which can be co-constructed between the researcher, the readers of the research who evaluate the researcher, and participants who are invited to comment on the results and have access to the research. They brought seven factors as the requirements for establishing the warrantability of research: 1- orderliness or documentation, 2- accountability, 3- demonstration, 4- patterns, 5- coherence, 6- plausibility, and 7-

fruitfulness or the implication(s) of research, generating novel explanation in terms for future works (Wood & Kroger, 2000 as cited by Saldanha & O'Brien, 2013, pp. 41-40)

At the end of chapter two, they explained the importance of regarding research ethics and these eight topics, 1- ethics in context (regarding the context that the research is conducted and can be preserved by asking the question of why the researcher wants to study the particular selected topic and does this research help the people of the particular context that the researcher surveys them in his/her research, 2- ethical approval, 3- informed consent, 4- rejecting deception that means refusing to cheating participants that for example this research is for the defined topic but in fact, the research is for a different or a related topic in order to ensure that the participants' behavior don't change when they become aware that what aspect is being researched or studied, 5- controlling and blocking the effect of power relations on participants, 6- protection from harm (physically and psychologically), 7- controlling the conditions of any internet-mediated research by considering four kinds of participants: a- recruited and identifiable, b- unaware and identifiable, c- recruited and anonymous, and d- unaware and anonymous participants, and regarding research ethics by considering these two main question of does the participants aware that they are being studied, and are they identifiable or not? (Saldanha & O'Brien, 2013, p. 47), and 8- plagiarism that may be accidental or un-accidental were mentioned and foregrounded by them (Saldanha & O'Brien, 2013, pp. 41-49).

Chapter 3 related to the product-oriented research and started with explaining descriptive, and explanatory approaches and introduced Critical Discourse Analysis (CDA) and Corpus Linguistics (CL). They expressed that CDA is brought along with the name of Norman Fairclough and distinguished it from Discourse analysis (DA) by focusing on describing discourse, critical, and analysis. They stated that "critical discourse analysis goes a step further than DA in its understanding of discourse as language in context" (Saldanha & O'Brien, 2013, p. 53). They pointed to the relationship between the text and "the order of discourse" (Saldanha & O'Brien, 2013, p. 53), and called it as a socially and ideologically practice that governs the minds of people and influences their speech and talk, behavior, and their practice and imposes inequalities. Next, they described corpus linguistics as the study of language based on the corpus, or corpora. Corpus means a set of related and structured texts that has been gathered and corpora means corpus in a plural sense. It is important to note the critics over CDA as its stress on the selection of text that must be significance or typical based on it and the typicality must be carefully justified, for it may on the basis of subjectivity and consequently leads to three problems as the research may be based on the privileged knowledge or the impossibility of reproducing or replicating the results, and conclusion of the research, or the text be selected due to the preconceived ideas which leads to producing the risk of circular argument.

They brought the expression of Baker (2006, as cited by Saldanha & O'Brien, 2013) that considering corpora for any CDA research can almost but not completely remove the risk of subjectivity like being bias and generalizing results, since it gives us the chance of studying the incremental or cumulative effect of discourse. After, four different types of

designing studies were described by them including, corpus-based design, corpus-driven design, problem-based design, argument-centered design. The difference between the first two designs is the first approach begins with considering a pre-existing theory as a proper tool for a corpus data, while the second one, step by step leads to producing a new theory, or hypothesis or to challenge the pre-formulated descriptions and theories.

Faircloughian CDA framework was regarded as Fairclough's problem-based approach which is started with identifying a social problem embedded within the linguistic aspects, such as capitalism or hegemony of a particular society which causes producing inequalities and discriminations. Moreover, on the basis of their explanations, CDA often is used to propose a theoretical argument by considering the carefully selected evidences and proofs that are extracted from texts and so, as a design of study it is argument-centered design which is the complex of both conceptual study and empirical study. Furthermore, they emphasized on this point that as the minimum the circumstances and the context of production of discourse, the circumstances and the context of recipients of translation, bio-social information such as their gender and professional background about the agents that are translators, or interpreters along with gathering basic and fundamental information about the other agents such as editors, publishers, or the status of officials in community interpreting.

After, monolingual comparable corpora, was explained as it contains two sets of monolingual texts. One set is regarded as the original texts, and another set is translations in such a way that both of them are in the same language, in order to explore and reveal linguistic patterns as the distinctive of translations which are in opposition to the related elements of the original texts which are produced in a non-translational context or environment. The noted that, parallel corpora may be bi-directional. It means that in addition to contain two sets of texts with two different languages as original texts (language A) and their translations (language B), it may contain source texts of language B and their translations into language A, that as an example, they pointed to COMPARA as the bilingual parallel corpora. They also noted that, parallel corpora can be the translations of the same source texts into several different languages, and they brought the case of MULTEXT-East "1984" corpus, that includes the novel of Orwell entitled, "1984" and its translations into Romanian, Slovene, Bulgarian, Czech, Estonian, and Hungarian. It is important to point to this fact that a direct translational relationship does not have to exist, when the researcher uses multilingual parallel corpora.

In fact, Multilingual parallel corpora can be considered as it is shaped by several bilingual parallel texts, or it contains the translations of a particular source text, in different languages, from two to more languages. It also should be noted that, parallel corpus or parallel corpora due to language combinations have been used based on the focus on norms rather than linguistic universals, such as Kenny's (2001, as cited by Saldanha & O'Brien, 2013) study of normalization in the translation that was performed from German to English. The other usage of parallel corpora are the investigations of stylistic issues in translations, investigating the features of translated text. They stated that the combination of comparable and parallel corpora can provide more valuable insights, for example the effects of source text and language, and genre-specific patterns on

translation. In addition, learner translation corpora, is used for researches related to translation teaching and learning and to examine the development of translation competence. As an example of this corpora Learner translation corpus (LTC) which is produced by MeLLANGE project. They pointed to the main challenge and difficulty that relates to providing a learner translation corpus where all variables, such as direction of translation, type of translation problem and teaching method, level of proficiency, that can affect the output of learners and consequently their output can vary, can be controlled.

Next, different kinds of annotation were brought as documentary annotation, structural annotation, and linguistic annotation. Documentary annotation is for transparency and replicability that includes extra-textual information such as the details about the author, date of publication, etc., and also information about the text processing, whether its proof-read, or scanned, etc., which is in the form of header that it is a separate file but associated with the text. Structural annotation includes paragraph breaks, headings, subdivisions, footnotes, and the like. They also foregrounded Text Encoding Initiative (TEI) for providing guidelines to encode structural information and documentary information and to prepare them on the basis of the principles of the standards of representation as the electronic texts that can be used for research and teaching. Linguistic annotation includes different kinds and categories such as syntactic, semantic, discourse, lexical, phonetic, and prosodic annotation.

The process of linguistic annotating is also called tagging. The most common use of tagging is Part of Speech Tagging (POS) that can be done by the computer programs that are called taggers. POS-tagged parallel bidirectional corpus that contains German and English languages was used by the scholar, Teich (2003, as cited by Saldanha & O'Brien, 2013) to identify and find the normalization and interference in translations to answer such this designed hypothesis that the number of post-modification instances by the act of translation from German to English due to normalization will be decreased in comparison with German original texts, by getting information from pre-modification, post-modification, word classes, and usage of passive, predicative, adjective and attributive.

They mentioned that specialized corpora can be annotated in order to serve particular aims. For example, parts of the translation learner corpora can be annotated on the basis of considering the hierarchical scheme of error-typology, that contains content-related and language related errors, for translation errors. Thus, linguistic annotation is considered as the enrichment of raw or plain texts (unannotated texts). Additionally, in tagged texts or pursued corpus, the categories of analysis are not seen as they are derived from data, but they are imposed by the researcher based on regarding the pre-existing theory. In other words, something that can challenge the pre-existing theory cannot be revealed and so, a restriction is imposed on findings. The other term that explained by them is alignment that means connecting or associating the unit of source text with the corresponding target text as a parallel concordancer. The linguistic toolkits or analytical concepts to analyze texts and corpora relate to the context of culture of production which

its related concepts are genre, register, and the context of situation that contains three dimensions of field, tenor, and mode.

The next was mentioning Halliday's Systemic Functional Grammar (SFG) that foregrounded ideational function, interpersonal function, and textual function, and also the concepts of transitivity that considered the participants and circumstances, and also the concept of pragmatics that emphasizes on producing and understanding the meaning of language in relation to the context of production. Face was another key term of this book that described as two different kinds negative face, and positive face that indicated to the freedom of action and freedom from imposition, and positive face that means a public self-image that language users or participants have desire to be approved or appreciated by the others. After that, three relations of Faircloughian CDA were brought as external relations that includes considering the factors of social structures, social practices, and social events, internal relations that includes considering the factors of discourse, genre, and style, and internal relations that includes grammar, vocabulary, and the areas of phonology and graphology. Next, corpus- assisted CDA's stages that Baker (2008, as cite by Saldanha & O'Brien, 2013, pp. 85- 86) was suggested were brought as these followings:

- 1-Context-based analysis of topic via history/politics/culture/etymology. Identify existing topoi/discourses/strategies via wider reading, reference to other studies using critical discourse analysis;
- 2-Establish research questions/corpus building procedures;
- 3-Corpus analysis of frequencies, clusters, keywords, dispersion, etc. – identify potential sites of interest in the corpus along with possible discourses/topoi/strategies, relate to those existing in the literature;
- 4-Qualitative or CDA analysis of a smaller, representative set of data (e.g. concordances of certain lexical items or of a particular text or set of texts within the corpus) – identify discourses/topoi/strategies;
- 5-Formulation of new hypotheses or research questions;
- 6-Further corpus analysis based on new hypotheses, identify further discourses/topoi/strategies, etc.;
- 7-Analysis of intertextuality or interdiscursivity based on findings from corpus analysis;
- 8-New hypotheses;
- 9-Further corpus analysis, identify additional discourses/topoi/strategies, etc.

After, some applications and tools were introduced that are used for managing, analyzing, and querying the annotated corpus or corpora including, XAIRA, the MODNLP- tec suite, and the Corpus Workbench (CBW). They further explained that the most analyses are on the basis of word indexes that the information can be shown as in the form of frequency lists of word forms or types. the main problem of these tools perhaps is that of ignoring the features of words when they be homographs and also the problem of spelling variant. They noted that, lemmatization can solve the problem of the types of words by labeling,

for example, “am” and “were” as the forms or types of the lemma, “be”, and also POS tagging. They also stated that, word indexes provided basic and fundamental statistics such as type-token ratio, and average length of sentence. The result of the type-token ratio can indicate to the lexical specificity and complicity in a text or corpus. The defined type-token ratio as the result of the ratio of the number of different words (types/forms) to the number of running words as tokens in a text. They also described the meaning of this term, keywords as the items of the text that their frequency is significantly high or low in comparison with regarding the frequency in another corpus or text. They clarified that low frequency relates to the case of negative keywords. Lexical density was another term that means the ratio that is obtained from the number of lexical words to the number of tokens or running words in a text or corpus.

They also said that, word indexes or analyses of frequency, also gives the information of dispersion as identifying does a word has a regular occurrence across the whole text or mostly appears in a small section, or more clearly on where the word occurs in a text or corpus. The data that are produced by such these above applications that show the statistical frequencies of word, are illustrated by their co-text. Co-text can be comprehended when the analyzer looks at the specified node with its linguistic environment or context, that the word is surrounded by them. In other words, the words of the right and left of a node are recognized as its co-text. So, by looking at the concordance that was defined as a list of all occurrences of a specified word or node with its linguistic environment or immediate context, or co-text, the researcher or analyzer can see the specified lexical item in its defined co-texts. Concordance is very useful for exploring and identifying the collocations, semantic preference, colligations, and semantic prosodies. Regular expressions is another way of looking at a particular node or word, where a wildcards are used to offer a match between fore example “colou?r” and “color” and “colour”. Regular expressions are useful for more complex searches and for providing more flexibility. Semantic prosody, is another way of looking at a particular word type as a word or the collocates of a word, that occurs in a linguistic environment or its co-text, mostly is used in a neutral, negative, or positive way, for showing difficulty like trying or struggling, or the other ways. Semantic preference can be seen by the relationship between a word-type or form or a lemma and “a set of semantically related words” (Saldanha & O’Brien, 2013, p. 91).

For example, like the relation between the word, make sense, as a lemma or word-type, and these words of made sense, making sense, makes sense or the word, “source” that can show a preference for these words of cite, quote, claim, confirm, say and report. Next, translation quality assessment was elaborated and some related key terms such as evaluation, error typology, comprehensibility, readability, usability, acceptability, adequacy, fluency, mismatches, ranking, and quality were explained. Two important key terms are the definitions that are related to the two kinds of mismatches as dimensional mismatches and non-dimensional mismatches. Dimensional mismatches in translation quality assessment, refer to the pragmatic errors that two factors of language users and language use can lead such these errors to happen. On the other hand, non-dimensional mismatches are recognized when breaches of target language system happen and/or

mismatches or errors at the level of the denotative meanings of original text and translation elements happen.

Furthermore, the work of Automatic Evaluation Metrics (AEMs), such as bi-lingual evaluation understudy (BLEU), as the assessors of the quality of the translations that are performed by the machine translations were described as comparing the machine translated sentences with one or many gold standard or reference sentences that are performed by a human translator and scoring them on the basis of the extent of similarity to the translations of human translator. The main problem of AEMs is that for example, how the score of 0.53 means on an AEM scale? But, this criticism was tackled by the community of machine translation research by directing directing a serious attention to research into correlations between human quality assessment and AEMs. Two statistics that are used for measuring the inter-rater reliability or inter-coder agreement are Cohen's Kappa coefficient (k) and Fleiss' Kappa coefficient for measuring agreement between more than two coders or raters. If agreement, between coders be complete K is equal with 1 and if it not Kappa score is equal with 0. It should be noted that, the closer score to one is defined as the higher level of agreement. They are considered as more robust measures in comparison with the simple percent agreement calculation.

They also foregrounded the statement of Bayel and Paul (2011, as cited by Saldanha & O'Brien, 2013) over the necessity of participating evaluators in the process of assessing the quality of translation who be viewed as with the same degree of expertise, along with recommending intensive training, and emphasizing on the number of them as what defined by Bayel and Paul (2011, as cited by Saldanha & O'Brien, 2013), the use of five or more evaluators for very critical tasks and three or four for less critical tasks, where not very high quality be critical. As said before, chapter four was allotted to elaborating the process-oriented research. They explained it as a seeking for understanding the behavior, expertise, and competence of interpreter or translator, the cognitive processes or mental operations that direct and control this act, the relations between interpreted or translated product and cognition, and also understanding the impact of social context on the process of translation. They explained different ways of knowing the mental operations of translator and his/her cognitive efforts in the process of translation. One of these ways is the EU-funded Casmacat project that is founded on the effort of seeking a way for building a cognitive model when translator or interpreter is in the process of his/her translation, by using computer tools.

The important key term of this chapter was metacognition that means the knowledge of interpreter or translator about the process of his/her translation, such as how he/she solved the particular problem, or managed the uncertainty, and what strategies he/she used. After, the model of working memory (WM) capacity was mentioned and foregrounded as the item that play a prominent role and the researchers of this area are interested in establishing their researches on this subject that how brain of the interpreter or translator the process of translation which is a bilingual process. Cognitive rhythm was defined as the explosion creativity between stops or pauses that can be recorded and measured and analyzed by the translation process researchers and also can show that how the process of translation is separated into different phases of orientation,

drafting, revision, etc., and the time that was put on various activities such as typing, insertions, deletions, revising, the use of information resources like online resources and reference works.

The other key term was experts that defined by them as those who are trained and are considered as bilingual professional translators. On the other hand, non-expert translators are those who are considered by them as novices or those bilinguals who are at the early stages or those who are not trained. semi-experts are those who are under training whether undergraduate student or postgraduate students or those who are under the rigorous translator training to start their career as translator. Next, some introspective methods were introduced by them such as think-aloud protocol (TAP), keystroke logging software, eye tracking, and screen recording, scan brain activity, like Functional Magnetic Resonance Imaging (fMRI). They emphasized that, all process-oriented researches even by getting helps from these above-mentioned tools are essentially regarded as an indirect measurement of translation processes which is valuable and valid. It should be noted that in process-oriented researches these two different designs, between-subject design, within-subject design, can be constructed. Between-subject design is defined as design of research when individuals or participants are categorized into two different groups and the results be compared with each other.

For example, the process-oriented research be conducted between the two groups of professional translator and semi-professional translator. On the other hand, within-subject design defined as when the research be conducted on the basis of analyzing the process of the translation of one participant or individual under different task conditions. They noted that, within-subject design reduces the number of participants to one as what that the research benefits from its feature, but, the carry-over effect as the most important and one of the drawbacks, threats this design, where repeating a task has an impact on the latter tasks that are taken into account as a sequence of tasks. They stressed on the randomization of task sequence to control the impact of it, along with emphasizing on the objectives of a research to rejecting the assumption that may be formed in the mind of the readers as showing its merit to abandonment usage of it.

They further stated that, the main challenge in the process-oriented analyses is how to define and identify the routine researches in the area of translation, and they emphasized on the directionality like, L1 to L2 or vice versa, as operational parameter to make distinct the routine research from non-routine researches. After, six factors of characterizing data that are elicited in the area of process-oriented research were brought as these followings:

- Number of Participants,
- Research context ('when', 'where', and 'how'),
- Recently of the event,
- Mode of elicitation and response (oral vs. written instructions),
- Formality of elicitation,
- Degree of external intervention (Saldanha & O'Brien, 2013, pp. 116).

They also stated that, if such these elements be implemented in many researches that relate to process-oriented translation studies, as a high-level basis, they can provide the right condition for cross-study comparison. They pointed to multi-factorial or cross-factorial research designs when more than one factor or variable and more than one condition be taken into account by the researcher like the two parameters of context and expertise be regarded and the impact of them on the creativity of translator be investigated. So, by these factors four conditions can be designed as transcoding text, translating text, bilingual transcoding text, and bilingual translating text. They foregrounded the act of randomization in such this design to prevent the carry-over effect and the other possible effects that can be due to the sequence that tasks are preformed or texts are processed. After, they studied the probable effects that can threaten the validity and reliability of the process-oriented research such as whitecoat effect that threatens the participants' behavior and may happen, when the environment be too library-like or the participants face an unfamiliar environment. For example, collecting data from students in a class even the class be not a formal class or unfamiliar tools like strange keyboard, monitor and software can cause such this effect on the result.

Longitudinal studies must be employed for comparing novice translators with expert translators and interviews and diary of expert translator as the supplementary data along with tests of expert performance can help such this study. Quasi-longitudinal design also can be emplaced of a longitudinal research and instead of selecting certain particular participants, who must be under the study for many years, for different stages of the development of study different participants be selected and the gathered data be compared. For example, first-year and third-year under-graduate students be studied and then graduate students and post-graduate students be studied, and finally the gathered data be compared with each other to reach a rational result. They noted that, pilot stages and testing the tools and equipment of data gathering like computers are very necessary and significant, to prevent any failing in this area or to minimize the risks of it, when the process-oriented research is structured.

Next, they enumerated the types of the units of analysis in process-oriented translation researches, as problem indicator, the attention unit, and temporal unit. Problem indicator relates to the models and frameworks that are proposed for recognition of problem, solution proposal and problem resolution. It is important to note that strategy is a problem solution procedure. Attention units defined as the hierarchical segments of attention like clause that as a smaller unit is inserted into the large unit of text, or as what is defined by Alves and vale (2009, 2011, as cited by Saldanha & O'Brien, 2013) as micro units and macro units, but they are not referred them to as the linguistic units. They defined them "as the flow of continuous target text production separated by pauses" (Saldanha & O'Brien, 2013, pp. 120). Pause lengths are arbitrary and different scholars selected and proposed different length of pause for determining the unit of analyses. For example, long pauses are those that may relate to or be due to the cognitive processing of translator in the process of translation, while short pauses may be caused by the physiological constrains like the low speed of the participant or translator in typing. So,

temporal unit of analysis can be obtained by pauses and they are pauses that separate process indicators on the basis of their length of time.

For example, pauses that determine the temporal units of production, deletions, and insertions. Each researcher on the basis of the research design, and the objective(s) of his/her research can select the most appropriate unit of analysis. It should be noted that when the researcher uses triangulation method for gathering data, the unit of analysis may not be matched with each other. For example, the pause in verbal reporting cannot be not coincide with the pause that related to typing that was shown by using keystroke logging. In other words, in triangulation, the problem of segmentation must be considered. So, they concluded that triangulation come with important methodological challenges.

After, they described they key term, introspection, as the process of self-observing and reporting, which happens as TAP, or as written protocols, retrospectively getting report. For retrospectively getting report from the participant of the research as the translator, the records or documents of his/her work like getting the screenshot of the task is essentially needed for such this verbal report. Such this document or record that is provided for getting retrospectively verbal reports leads to the outcome that is named as a cued retrospective protocol. It should be noted that, in written protocol the participant or translator writes his/her experience currently or simultaneously along with doing the task or retrospectively. The case that was named as "integrated problem and decision report" or IPDR for short, can be a good example of written introspection (as stated by Gaile, 2004; Pavlovic, 2009, as cited by Saldanha & O'Brien, 2013, p. 123).

The weakness of the introspection is the cognitive effect, for example the effect of thinking aloud of participant, on the very processes that the aim of research is to investigate or what that can be interpreted as the cognitive processes. Additionally, they expressed that TAP does not change the structure of task process but it can slow down it. They also stated that, although think-aloud protocol is good for generating hypothesis, it is not good for testing hypothesis. They expressed that, TAP has the weakness of incompleteness, since some behaviors and cognitive processes like habitual actions, can be frequently or automatically occurred without any reportable thoughts as verbal reports. In other words, these actions are unavailable for verbalization. Furthermore, those translations with socially agreed performance characteristics or those that are performed based on translator's consultations with others through TAP, imposed constraints on translator's functional mental representations and successful thinking. They added, it happened, since translation does not have "the luxury of socially agreed performance characteristics" (Saldanha & O'Brien, 2013, p. 123). They also enumerated the other weaknesses of TAP such lack of systematicity, the possibility of distracting translator, inconsistency of verbal reports with observed behaviors. Furthermore, they stressed on this point that, concurrent TAP for those researches that are designed on the basis of considering translation task time as the unit of measurement or temporal unit cannot be suggested, because of its slow-down effect on data elicitation, and retrospective verbalization can be more suitable for such this research design.

But, TAP, still have been considered by the scholars of TS as a valid data elicitation method for translation process-oriented researches, such as examining translator's

competency, how professional translators work, and producing models of translation process for training translators. They also said that, TAP can be regarded as a good tool for handling translation problem, since it lets the participants to carefully think about the faced translation problem in the process of translating and to produce a resolution for solving it and so, their performance is mostly improved, and leads to task improvement or better translation. They also emphasized on the result of the comparison of participants who are divided into two groups of those who done their task with thinking aloud and those who completed their task silently, that those who thought-aloud were more systematic and also successful problem solvers. They foregrounded the slow-down effect which mostly happens during dialogue or conversation protocols when two participants discuss about their shared task. They bolded the advantage of written report in comparison with verbal report as it does need to be re-encoded, while its disadvantage as if written report be occurred after completing the task of translation is that it can be affected by memory limitation of participant or translator. They noted that, verbal reporting is known as an uncommon activity by the most of people and can affect the task performance.

They explained that the use of verbal reporting in terms of considering its advantages and disadvantages must be based on the aim of study and must depend on the variables that are selected to be examined. They bolded the influence of culture that verbalization can hamper the acts of participants in the task of translation, by bringing this example that the study of Clemmensen et al (2008, as cited by Saldanha & O'Brien, 2013) proved that thinking aloud hampered the acts of eastern people like Chinese who are heavily affected by their culture, while silence working hampered western people like people of United States of America and Europeans. Verbalization or TAP in its nature, in the task of translation, is a bilingual protocol, and if the participant chooses the language of verbalization, so the process of annotation, coding, and analysis will be more complex, than when the participant uses only two languages, as are used by him/her in the task of translation and also for verbalization. They said that, verbal reports must be transcribed before starting the process of analyzing.

Next, they enumerated some numbers of computer programs that do the task of transcription such as Express Scribe, Hyper Transcribe, Transcriber, Voicewalker, and Transana. They explained the weaknesses of transcription programs such as due to having their own standards they may lead to change the spelling of lexical variants, grammars, or punctuations. They introduced denaturalized transcription program that is created to retain the feature of oral reporting such as pauses, filler words, false starts, prosody. They also said that, transcription threats oral speech in such a way that it contains the risk of being unsuccessful in paying calling attention to the possible change of meaning in the conversion from oral speech to writing. They distinguished it from naturalized transcription that discursive features of oral representation of cognitive process are deleted and the result is more like reading the written language.

After, the act of coding was clarified by primarily pointing to the fact that the processes that totally form the process of translation cannot be directly coded from the act of verbal reporting or verbalization. So, it must firstly be transcribed. Coding can be happened at

the levels of micro and macro. Macro-level coding connects to episode. In here, solution-step analysis was defined as shedding the lights on the episode. Episode means a collapse view of the specific steps that are taken in the task of translating. They clarified their words by stating that macro-level coding and analysis are suitable for comparing general approaches to a problem or task. Micro-level coding deals with micro processes which makes each step. This below table can clearly show coding:

Table 1. Example of coded verbal protocol for making coffee (Experts think alouds, n.d.)

	Verbalization	action	code
1	OK, ready for breakfast – first gotta make coffee		
2	Let's see	[opens lid]	
3	... hmm, there's still the coffee grounds in here from yesterday, so let's throw those away...	[removes filter tray and dumps old grounds into trash]	1.1. Remove grounds
4	Alright, lets get the filters	[opens cabinet and takes filter out of bag]	2.1.1. Get filter
	[...]		
13	Not there	[Opens dishwasher]	1.2 Clean pot -- must've been done earlier
	[...]		

And creating a hierarchical task analysis like this below figure can be done:

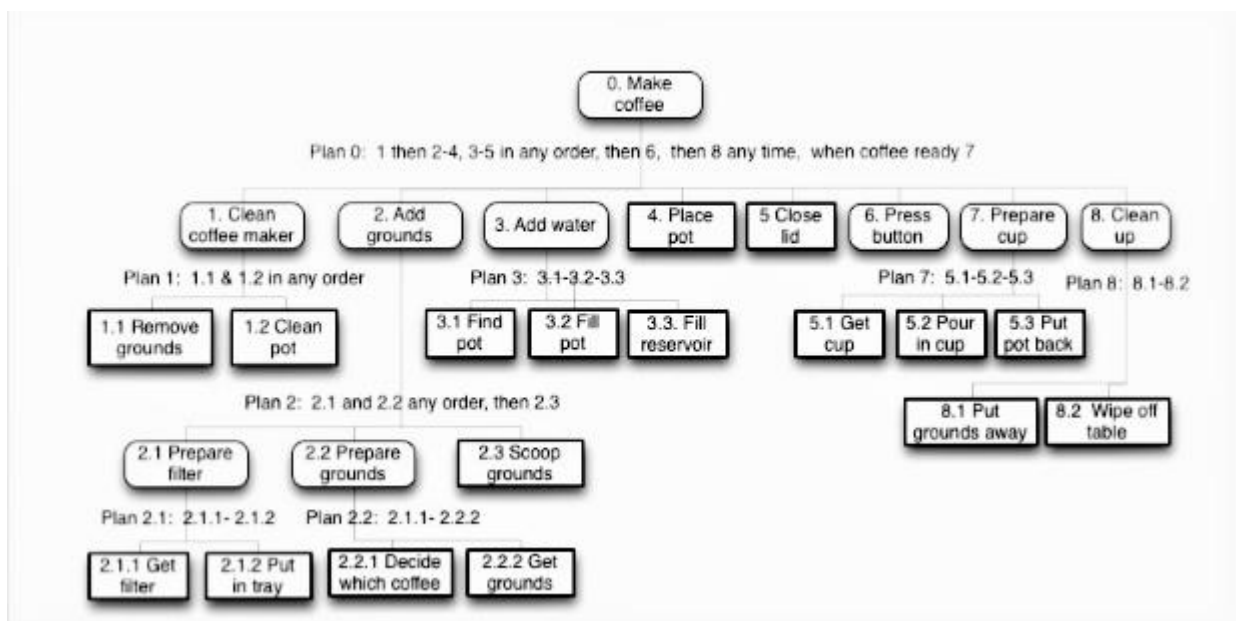


Figure 1. Hierarchical task analysis of the verbal protocol in Table 1 (Experts think alouds, n.d.)

They explained that triangulation causes the validity and reliability of research, due to not matching the units of analysis, to be discussed and challenged, so they suggested developing a coding manual as a guide to know how to categorize the data, and for overcoming the challenge of not matching data segmentations. The next was emphasizing on Inter-coder reliability for assessing the decision of coders on allocating a particular category to specific data. Inter-coder reliability, as a measure was determined by the presents of fifty as indicating to the low reliability and ninety-five as having the high degree of reliability. They stated that, the common measures for inter-coder agreement are Cohen's Kappa and Fleiss Kappa. The first one is used for two coders, while the second one is for measuring the agreement of more than two coders. They also enumerated the some emphasizes, and proposals of scholars such as stressing on standardization of coding schemes, using asset management system to access the data of translation process on the internet, and text annotation system that works on the basis of the Text Encoding Initiative (TEI) for translation process protocols.

After, Keystroke logging was described as the special software program the enables the researchers to see the recorded all key pressed on the keyboard such as navigational characters, punctuation symbols, and text production characters, pauses between key presses, and mouse movement, and then two software programs, Inputlog or BB FlashBack, and Translog, as the important keystroke logging tools were brought. They bolded the advantage of keystroke log tools as they in comparison with TAP don't slow down the task, nor they act as the interferer by putting extra demands on translator's cognitive resources. They recommended the usage of this tool with at least one other data elicitation method like eye tracking, or verbal reporting to increase the researcher's ability of making inferences about the processes of translating. They brought the weakness of keystroke logging tools that the line between the record of screen activity and keystroke activity is fuzzy. They stated that keystroke logging tools record in two aspects of micro and macro levels. Macro level indicates to recording total time of translation, and showing the time that spend for orientation, drafting, and revision, while micro level shows pauses, frequency of different kinds of pauses, duration, positioning, and other that represent the cognitive efforts, the units of task production and the length of them, and their numbers, and also the numbers, essence, and time of revision activities that can be regarded as cognitive efforts.

They enumerated the advantages and weaknesses of using Translog tool as it shows in parallel the activity of translation and corresponding original text in a dynamic linear manner, and and its advantages over the use of verbal reporting as it does not slow-down the task of translation and also does not put the extra demand on participants' cognitive resources, while its disadvantages were enumerated as it isolates the participants in its environment in such a way that any activity out of its environment like searching in internet cannot be recorded that curtails the usual translation routines, its weakness regarded to not having the spell-checking or formatting features, and also mentioned its inability of interacting with other translation technologies like translation memory tools. On the other hand, they bolded the advantage of Inpulog and BB FlashBack as they provide the condition of independently recording the keystroke and giving the flexibility

of using the other software programs. After, the key term, eye tracking, was defined as the process of recording the movements of eyes from one point to another and the stare of the participant or translator in the process of translating when the retinas stabilize on the object. They clarified the origin of eye tracking as eye movement hypothesis proposed by Just and Carpenter in 1980 that indicates to this point that the process of gazing or fixation and brain's process are both happen in relation to each other and in a same time or with no time lagging. Fixations determine the points of the focus of attention and cognitive efforts.

Eye tracking also can visualize the path of gazing or the sequences of fixations on an object or stimulus during a period of time. They also stated that, voice, light, heavy make-up, and caffeine can affect the size of retinas and the results of eye tracking. They further explained that, some participant may not be suited for eye tracking and their power must firstly be considered for example, assessing them by calibration or when the research needs in minimum 20 participants, instead of this number 25 be regarded. They also foregrounded the importance of relaxing during the process of calibration that prevent participants moving their head and changing the position of head, although the modern tool can compensate small variation of head position, along with using non-swing chair and being adjusted in height, and using chin rest as facilitators. But this artificial environment is criticized and challenged for limiting the natural acts of translator, or for ecologically valid translation production tasks is highly questioned by the researchers. They emphasized on fixation count, average fixation duration, and observation length as are the typical in eye tracking studies. They stated that eye tracking covers short text, and that is what the researcher faces the dilemma of ecological validity.

They added that, eye tracker does not able to show 100 percent accurately fixation of eyes on a particular word, phrase, or sentence; cannot completely shows the focus of attention. They stated that for showing the exact fixation or increasing the reliability, nowadays the researchers use the font size of 14 to 16 along with double spaces, paying attention to this point that the text not goes to very right or left of the monitor. Additionally, they explained gaze reply as the part of the work of eye tracker that retrospectively shows by a colored dot where the translator stared or looked at any point in the process of translating. Moreover, they state that, the attentional data are represented in the form of heatmaps or hotspots in such a way that statically showed the part of receiving the most attention by red color, and then other parts of interface by fore example yellow and then green that indicate to decreasing the attention in comparison with the part of interface that was marked by red color as it is received the most attention. They emphasized on using linear mixed effect modeling (LMER) that within the approach known as multiple regression analysis, instead of using Anova, in a naturalistic environment of translation, since Anova has the high degree of control over all variables, and multiple regression technique covers and manages with the lack of control on variables.

They enumerated some other methods that can act as complementary to or alternatives for including contextual enquiry that was defined as interviewing with participants in their natural environment during their engaging in the task of translation and it can be protected by theory known as situated embodied cognition. Personality profiling, as the

need of describing the profile of participants for screening their competence and also in order to become aware of the impacts of their personality traits such as attitude, experience, personality on translation competence and the used strategies. They suggested psychometric tests that was defined as the forms of standard questionnaires to measure personality types or particular traits, and psychological measurements such as electroencephalography (EGG) galvanic skin test, and functional magnetic resonance imaging (fMRI).

After, chapter five started with these two key terms, respondent as participants who answer to the questions of questionnaire and informant as the partner in research or study who is more active in research and in structuring the knowledge that is to be regarded as the result of the research. Questionnaire was foregrounded by them as the research instrument and defined as "(1) a list of questions each with a range of answers; (2) a format that enables standardized, relatively structured, data to be gathered about each of a (usually) large number of cases" (Matthews & Ross, 2010, p. 201, as cited by Saldanha & O'Brien, 2013, p. 151). They enumerated the usages of questionnaire, as it can be used for collecting background information of participants and/or for collecting data related to attitudes, opinions, facts, behaviors, etc. Most usages of questionnaire are in the areas of translation profession, surveying translation students' attitudes and opinions about the subjects of learning and teaching, and technologies.

They are popular, since they are formed based on a structured data collection that in comparison with interview takes less time of surveying and analyzing data, and so they can be used for a large scale of population and in terms of assuming sampled population generalization of the outcomes of the analyses of the questionnaires can be generalized. On the other hand, weaknesses of questionnaires as it is easy to get the design and administration of a wrong questionnaire, and also they cannot be regarded as the best instrument for collecting the exploratory data like personal experiences, emotions, and opinions, and so more in-depth interviews must be added, along with Hawthorne effect, the most challenge, the risk of low response rate due to participants' feeling like a/some constraint(s) on their response, nonresponse error, sampling error when some of the population in comparison with other parts of population have high degree of being included in a research, and measurement error when the real response is different from the fact or true response of participants like their beliefs, and facts which can be due to sensitivity, or ambiguity of the question or due to regarding bias effects of the researcher.

For designing a questionnaire, they emphasized on the first determinant, operationalizing the construct that the aim of your study is based on it, for example the aim be regarded as knowing about the perceived status of translators in a particular country, first of all the status as it is a subjective, context-dependent, and abstract and cannot be regarded as an objective and absolute notion, must be defined and operationalized by determining its vital parameters, and then the questions must be design that clearly address each of its parameters. they added regarding two important elements of language and the short length of questionnaire before designing or discussing on its technical details. Different questionnaires take different time of responding to its questions, which totally as mentioned before must be as short as possible to avoid

participants' non-completion. It should be noted that there is not the ideal or definable number of questions for questionnaire.

They stressed on considering some vital elements to design the questionnaire such as, asking the right question, regarding and respecting target population's time that they can response to the questionnaire, regarding progress indicator for the participants, filter or allow the participants to skip some of them that does not relate to them or not applicable to them, commencing or starting your questionnaire with fact-finding, background questions that used for collecting personal data, paying attention to separating background questions from the other questions, because of considering ethical aspects, informing or alerting participant in advance about the sensitive questions, asking participants to pre-defined categories, since it affects their responses to questions. Your wording must be neutral and not leading the answer that the researcher may looking for, and the question must not build implicit assumptions. They stressed on this point, answering those questions that are based on hypothetical situations are not reliable and valuable, since such these answers are rather or more considered as socially-desirable or idealized responses. questionnaire can have two kinds of question, close and open. They explained close question as it "might ask the respondent to select one item from a list or all that apply, or may even ask them to rank an item in a particular order" (Saldanha & O'Brien, 2013, p. 157).

On the other hand, open question enables participant to write the answer in a text box that is designed for short answer. "Open question will allow participants to explain their choices for the closed questions, to add further opinions, or to highlight an opinion they hold and which is not addressed in the questionnaire" (Saldanha & O'Brien, 2013, p. 157). They stated that adding at the end of the questionnaire, the question of anything you (as the participant) wish to add a box?, gives the feeling of satisfaction to participant, although it might not contribute to answer the research question, and can defuse the possible frustration of participants. Disadvantage of close question is it limits the participants to the options of question and does not allow to express the nuance thoughts, and for open question, the interpretation of it is difficult, and answering takes time and so, often participants refuse to answer, or because they don't have a well-formed response. A very useful tool in questionnaire is Likert scale that gives a series of five to nine responses from strongly agree to strongly disagree, that has the advantage of preventing participants conveniently selecting the mid-point or neutral point.

After, pilot testing was described as the testing of questionnaire "with an appropriate sample (by which we mean respondents who are part of your sample population and not family, friends, or colleagues who may not meet the criteria for inclusion in the sample)" (Saldanha & O'Brien, 2013, p. 158). The further explained that pilot testing or piloting should examine variant aspects such as the required time to fill out the questionnaire, the usability of it, its clarity, etc. they said that the researcher can get the feedback of participant over the subject of the orders of question, since they must be on the basis of a logical sequence and the link between them must be clear for participants. They stated that questionnaire must be like a story and began with simple question, and be followed by middle stage and be ended with complex questions.

They added some quantitative methods to assess the reliability and validity of questionnaire including, split-half method as the correlation test that allows the researcher to know about the internal relation of questions and the consistency of questionnaire for measuring a single construct, Cronbach alpha measure as re-applying split-half method for second time, test-retest method for measuring temporal stability that means giving the same group of participants the total of questionnaire, and then giving the same questionnaire to the same group when a period of time be elapsed, Pearson product moment correlation or Pearson's r when parametric data be normally distributed and measurement scale in interval or ratio and indicates to the values of +1 as perfect correlation and -1 as perfect negative correlation, and -0.7 shows a positive strong correlation, while -0.4 indicates to a weak negative correlation, Spearman's ρ for non-parametric data or those data that are not normally distributed and/or the measurement scale is nominal or ordinal.

In other words, Pearson's r is used for two variables that in a continuous order, the change of one affects the other one or linear relationship, but Spearman's ρ is useful for those two variables that change together and measuring the relation of one variable to the other one or monotonic relationship and it is based on the ranked values of each variable rather than raw data. They explained that, test-retest method is problematic because in a period of time the opinions may change and also participants can remember the answers that they gave and repeat them, so at least three mounts time must be considered and the time must not be very long or be very short because of the mentioned issues which can be raised, although longer time can be better than short time, and the final parallel forms test as giving two questionnaire to the same group with different wordings of same construct, as two versions, on two different occasions, and if the correlation of scores be high the questionnaire can be seen as reliable, this can tackle the problem of memory effect that may happen in test-retest method. They noted that rephrasing must equivalently reflect the same construct. Next, construct validity was bolded as the most important type of validity for questionnaire design. Concurrent validity, face validity described as seeing the questionnaire related to the concept or construct that it purports to measure or assess, and convergent validity that was defined as what that can work and be effective like parallel forms test which seeks to construct the validity of responses by asking the same question with different wordings, e.g. in what year were you born, how old are you?

They added that, for internet-based researches, the researchers can refer to the code and standard of internet-based surveys and can see ESOMAR (online) or CASOR (online). They explained different ways of data collecting by questionnaire as in the presence or absence of the researcher that data collecting can be executed in different ways, such as when the researcher is present, the questionnaire can be used as the toll of interviewing or or participants may (not) prefer to interact with the researcher, while it disadvantage can be the effect of power relations on the participants' responses., but the research by having a non-verbal communication and controlling his/her tone of speech and also by providing the the open ended-questions can tackle or reduce the possible effects. They bolded the point of a trade-off between the amount of collected data and the accuracy of

data. They also emphasized on random sampling or probability-based sampling as what that the researcher can make reliable inferences from it, when questionnaire be used for data collecting. They added, for establishing a reliability, the researcher must ask this prominent question, how many participants does he/she need to be fairly confident of the results (confident level) and also fairly accurate (confident interval).

They stated that the balance or relation between the size of sample and the time and cost for gathering data is undeniable. They stated that, the questionnaire response rate is very important factor that must seriously be regarded by the researcher in sampling method specially for postal questionnaire that the rate can be notoriously low as 25%. Hager-Johnson (2009, p. 97 as cited by Saldanha & O'Brien, 2013, p. 166) proposed a list of principles to increase the response rate of postal questionnaire as what brought in bellow:

- Keep your questionnaire as short as possible
- Stick to a clear, relatively conservative layout
- Include a pre-paid envelope with postal surveys
- Send a reminder after one or two weeks have passed
- Give advance warning
- Offer a small incentive if possible

They stated that face-to-face interview questionnaire are the good suggestion for increasing response rate, but it limits to those who geographically are near and also takes time. After, they enumerated different types of interview including structure interview that defined as carefully preparing a series of questions which must be carefully asked with the same wording and closed questions are usually used, unstructured interviews that was explained as free-flowing process of interviewing in such a way that some guiding questions are used by the interviewers and open ended questions can be asked by the interviewers and the researcher or interviewer are allowed to improvise, and semi-structured interviews that was described as the mix of both structured and unstructured interviews in such a way that an interview schedule along with more open-ended questions is used and the interviewer can change the order of questions or produce new questions and usually is used for life-story interviews.

The next key word was the definition of focus group as the group of people who their numbers can be 6 to 10 and together around 1 to 2 hours discuss over what that is the centered of issues as focused under the heading of a mediator or facilitator rather than an interviewer. Such this event can be structure or unstructured in such a way that in structured discussion the role of mediator is more directive, while in unstructured discussion the role is initiative in such a way that the mediator starts a topic and lets the participants' discussion flow in whatever direction they want or choose. They clarified their words by bringing the result of a study that revealed the role of mediator as what brought in bellow:

- To clarify a comment from a participant;

- To link a comment to that of another participant;
- To allocate a turn to a participant if needed;
- To bring the discussion back to the point of signed language interpreting (Napier, 2011, p. 65 as cited by Saldanha & O'Brien, 2013, p. 173).

Action research project as the other key term was defined by them as a piece of research constructed to give the knowledge related to change and improvement of the situation of participants which can produce practical and theoretical knowledge about the situation, and involves participants and structures an attitude of growth and self-development. For example, using it to establish the research on the effect of modifications made to a translation workflow process in an institution of translation, or a training programme. They further explained that using focus group in an action research can be possible. Some errors that may threaten focus group are false consensus, group polarization that may respond in an exaggerated way in comparison with individual member normally would. Vignettes was another key term that was introduced as the story-like or short narrating participants and events (interviews/focus groups). Interview profile was the other key term that was explained as the account of participant's contribution with their own words in first person. They added that both vignettes and interview profile can lead to refreshing the memory of the researcher in the process of analyzing data.

They added that interview can be happened via face to face connections, telephone, or internet that the usable tools for internet-mediated interviews may be Voice Over Internet Protocols (VOIP) that can be assimilated to the work of telephone, chat, and conferencing software which totally can be regarded as synchronous interaction, while asynchronous interview can be happened via email. Netiquette can be taken into account as the conventions that build online interaction, and standards politeness related to working in this environment. They brought the words of Hewson et al (2003, as cited by Saldanha & O'Brien, 2013) that relates to advantages of email interview. They said that, "Compared with traditional interview methods, the email interview may be less spontaneous and flowing, but it allows participants to answer in their own time and at their own convenience, which may encourage more detailed and carefully considered answers" (Hewson et al, 2003, p. 45, as cited by Saldanha & O'Brien, 2013, p. 188).

They enumerated the possible risks that can threaten the qualitative researches, such as confirmation bias, nothing the exception or norm that can be due to not having strong expectation, elite bias, going native when the researcher be very close to participants in such a way that participants' independently thinking be not possible. After, they explained several qualitative coding approaches such as, thematic analysis as "[a] process of working with raw data to identify and interpret key ideas or themes" (Matthews & Ross, 2010, p. 373 as cited by Saldanha & O'Brien, 2013, pp. 189-190), qualitative content analysis that consists of initial coding, and second-level coding, and grounded theory that has three-level coding system including open coding, axial coding, and selective coding. In qualitative content analysis, coding gradually moves from artificial code labels to more abstract code labels.

They further explained that, these two levels of coding, initial and second level coding proposed by that was proposed by Dörnyei (2007, as cited by Saldanha & O'Brien, 2013), as one way of approaching this process. They said that, initial coding can be understood as transparent and clear labels that can describe the meaning of the code, while second-level coding by going beyond the previous level notes emerging the patterns, they brought this example, clustering similar categories and through a broader label, particular hierarchies be introduced. They added that, answer of participants to the open-ended questions may refer to more than one theme or idea, and multiple codes might be assigned- in thematic analysis- to one response.

They bolded the subject of saturation as it determines the size of sample in qualitative researches and means data satisfaction and also gives the knowledge of where to stop in analyzing qualitative data which indicates that "iterative process stops producing new topics, ideas, and categories" (Saldanha & O'Brien, 2013, p. 192) and this point acts as the clue of where that makes the researcher satisfy or assures him/her to decide what is sufficient and relevant to address the particular research question. So, it cuts or lays aside the other possible avenues of research as they are beyond the limited scope of the current research and gives the notice of they must be explored in the frame of new research(es) or project(s). After, quantitative data were explained and measures of central tendency were highlighted as what that they are frequently applied for questionnaire including median, mean, minimum, mode, and maximum values. They noted that, for some statistical tests like the test of significance, the shape of the data distribution around the central tendencies must importantly be taken into account. Normal distribution was described as when the data be symmetrically distributed around the mean point in a bell curve like what is illustrated in bellow:

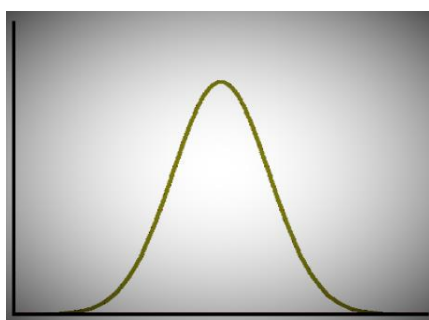


Figure 2. Normal distribution

They stated that, cumulative frequency tables, to describe data in more detail, can be used that values can be divided into four equal parts or quartile, five parts or quantiles, or even ten parts or deciles, and ranged from lowest to highest band. They clarified their words by bringing this example:

For example, if we are investigating the income of professional translators, we could rank our answers into the first 25% (or fourth quartile for lowest income), the next 25% is the third quartile, the next is the second quartile, and so on. In comparison with just reporting the mean value, this distribution into cumulative frequencies gives us a much

more detailed view of our data. We could then examine the inter-relation – a process known as cross-tabulation – between years of experience with income per quartile going from lowest to highest (Saldanha & O'Brien, 2013, p. 197).

They bolded reporting confidence levels due to not sampling the all population for the research that are the two types of confidence interval as the estimate margin of error and confidence level as reporting that how sure the researcher is that the result of his/her research is accurate. They said that confidence interval can be reported as minus or plus of x percent, e.g., within the range of +/-4% the researcher can say the result of the research is accurate. They also emphasized on 95% as its indication to the accurate result in survey research, while 85% shows that the researcher is lot less sure about that the result be accurate. They stated that, none of those mentioned parameters have not any indication to reliability of those results. They emphasized on using statistical test to ensure the reliability, such as chi square test, analysis of variance (ANOVA) for nominal data, paired t-test for ratio or interval data and for paired groups with normal distribution, Wilcoxon signed-rank, or Mann-Whitney U-tests for non-parametric data or those that are not normally distributed, Spearman's rank correlation rho for those data that are not normally distributed and the size of sample is smaller than 20, and Pearson's product moment correlation coefficient for data that are normally distributed.

In Pearson's product moment correlation coefficient and Spearman's rho, p-values indicate to the results that may be occurred by chance, e.g., in Pearson's product moment correlation coefficient, $r=0.90\%$ and $p=0.05$ that indicate to 95% confidence about the result, and 5% that the result may be occurred by chance. They noted that, 0 for both Pearson's r and Spearman's r indicates to the existence of no correlation, while +1 show highest and -1 indicates to high negative correlation. After, mixed methods research was explained as its typologies are based on the sequence of using methods and their dominance. By generally focusing on sequence, they enumerated three types of design including, a sequential approach, concurrent strategy (simultaneously using two methods), transformational strategy that can be quantizing data or less frequently qualitizing data, e.g., using the answers to questionnaires to develop a narrative profile of certain participants, as a way of developing a fuller understanding of a person and their responses (Saldanha & O'Brien, 2013, p. 203). At the end of this chapter, they added that, quantizing data can lead to reporting results in a precise manner. "For example, the frequency with which a specific theme occurs and its distribution or relationship with other variables" (Saldanha & O'Brien, 2013, p. 203).

Chapter 6 under the title of context-oriented research: case studies, started with defining the term, *case* that relates to the term, *case study* (Saldanha & O'Brien, 2013, p. 207):

Gillham (2000, p. 1) defines a case as a unit of human activity embedded in the real world; which can only be studied or understood in context; which exists in the here and now; that merges in with its context so that precise boundaries are difficult to draw. Yin (2009, p. 18), defines the case study as an empirical inquiry that investigates a contemporary phenomenon in depth and within its real-life context, especially when

the boundaries between the phenomenon and context are not clearly evident (as cited by Saldanha & O'Brien, 2013, p. 207).

They further explained that the case cannot never be regarded as a sample or an example, because case is considered as it is merit in its own and complete, and also case cannot put the assumption of using it to generalize the outcomes of research to the larger group that the particular case belongs to it. After, they stated emphasized on making distinction between case study and ethnography. They described case studies as a method where that the ethnographical approaches can be used. Next, they explained ethnography as “a holistic study of a particular culture, or community, and the aim is to attend to both the everyday details and routine text production as well as to the wider social context” (Koskinen, 2008, p. 37, as cited by Saldanha & O'Brien, 2013, p. 208). They stated that, this term also includes the emerged writing from the fieldwork data. Ethnography based on its old definition, means direct observing the behavioral activities of human in a particular context and within a particular culture to perceive the social realities from the view of those who live in the context and belong to the particular culture. They further explained that, the difference between the case studies and ethnography is more paradigmatic and theoretical than practical.

In other words, there is not a clear cut difference and it can be seen in focus; case studies, belongs to post-positivist and conservative tradition and the subjects of validity, reliability, and objectivity can be lost, while the experience of the researcher forms the research process in ethnography which belongs to postmodernism paradigm. The shared holistic philosophy as best understanding human behavior in the particular social context that needs rich gathering data and detailed data in an authentic manner was also bolded by them. After, the functions of case studies were enumerated as it can be useful for reaching the answer of the raised questions with why and how, for testing the workability of a theoretical framework, and for producing a hypothesis as oppose to hypothesis testing.

They added that, “case study can also challenge established theories and it may point to the need for a new theory in areas that have not received sufficient scholarly attention” (Saldanha & O'Brien, 2013, p. 210). Next, was defining different types of case, including single case or cross cases, that can be due to the homogenous or heterogeneous population of existing cases in sample case studies. They noted that, in heterogeneous cases, cross-case analysis is based on some degree of cross-unit (case) comparability which unit homogeneity must be taken into count. Embedded case study was the other type of case studies as taking into account the sub-units of analysis like several translations of a particular work, or some individuals that shape a group of translators and share specific features or characteristics, like a political agenda. They clarified their explanation by bringing the example of regarding all translations that censured or published in Nazi regime. They stated that, extreme cases, deviant cases, crucial or critical cases, and revelatory cases due to their unusual nature are selected as single case in case studies research. Extreme case is that the case in a particular aspect in comparison to the similar cases be extreme or striking; e.g., study of literary translation in Nazi regime. Deviant case is the case that in comparison with existed regularity in the similar cases it

has something unexpectedly or can be said, it unexpectedly happens. Revelatory case provides the opportunity of investigating a phenomenon that in previous was inaccessible to scientifically be researched. So, in revelatory case study, what that hitherto was unexplored is explored.

Crucial case or critical case is for testing a theory that when the case be very fit to the theory in the situation of one is very confident about the validity of theory or as the contrast of it be not well fit to the theory. On the other hand, typical case study was defined as the study of the case that is the representative of population, and so, can give use information about general population. For example, regrading translation in EU institutions as *a prototype case of an institutional setting*. They further explained:

Because of the looseness of the predictions and theories in the humanities and social sciences, one does not commonly find case studies that fit tightly into any of the categories described above (extreme, deviant, critical, revelatory or typical). However, the framework can still be useful because the principles underlying the classification are more general than a brief sketch of case types would lead us to think. [...]. The deductive logic of the 'risky' prediction may in fact be central to the case study enterprise. Whether a case study is convincing or not often rests on the reader's evaluation of how strong the evidence for an argument might be, and this in turn – wherever cross-case evidence is limited or no manipulated treatment can be devised – rests upon an estimation of the degree of 'fit' between a theory and the evidence at hand (Saldanha & O'Brien, 2013, p. 214).

They added that, sometimes two or more above mentioned categories may overlap in a case. For example, a case be typical and and also tells example. After, they explained types of limitation of case study including temporal boundaries that are self-explanatory, e.g., historical or contemporary situation or a period time be taken into account by the researcher, social boundaries, as taking it into account in terms of the type of group or population of case that shares some particular characteristics, such as what can be known as translation trainers, literary translators, or trainee translators, and spatial boundaries by its focus on geographical aspects such as particular region, or and nation. They noted that, seeing space can be also possible by considering it as (poly)system based on Even-Zohar, or field that is based on the terminology of Bourdieu, so, the study can be restricted to a particular genre that emerged within a particular nation, or a type of translation like audio-visual translation. They also bolded the the largeness of field and system and suggested a sort of artificial restriction that must be imposed based on rational and transparent factors. They emphasized on the aim of study and research question(s) as the two parameters in order to choose the source of information and limit them, since different sources can shed the light on the different aspects of the case of study, e.g., ethnographical observation and using official documents in the research of "translation in EU institutions as a prototype case of institutional setting" can fit to the research.

After, they enumerated some sources of data collecting in case study method as they bolded it for the all research, especially for literature review that in this step the subject area or the theoretical point of view of the researcher in this step must be contextualized,

including written sources, such as para-texts (blurb, footnote, preface, etc.), meta-texts (scientific articles, reviews, etc.), drafts, source text and its translation, verbal reports or interviews, observation be combined with video and/or audio recording that can be distinct into these two different types of observation, detached observation; as the researcher an outsider just observes the situation and participant observation; observing with having an active role in the group, and both of them can be threatened by Hawthorne effect or researcher as the inevitable observer. Detached observation is more structured than participant observation. Schedule listing the key events or behaviors, actions and also the aspects of particular behaviors that need to be focused on, and sampling those actions that frequently occurred. They said that, interval sampling can be used to check the situation; e.g., how many students are taking notes.

While, participant observation is very used in anthropological studies, and is a key parameter of ethnographic method, and physical artefacts, such as photographs, a work of art, technological devices like instruments or tools, and the like. They stated that, quantitative data can be extracted from those mentioned sources, specially from archival records, and also descriptive statistics is much more used than inferential statistics in case study. Inferential statistics is used in sample-case study to generalize the outcome to the population under the study, or to draw inferences from the data, while descriptive statistics quantitatively summarizes and describes data and allows to compare information. They noted that, inferential statistics does not for immediate meanings from the data, like likelihood that can be occurred due to chance or random. They noted that, the bridge between descriptive statistics to inferential statistics or sample to population is probability.

They suggest using a database to manage data. They added that these two types of notes must be recorded in research diary or log book, 1- personal notes, ideas that are regarded for the final reports, preliminary hypothesis, and the like, and 2- notes like remarks in the process of observing, expressions that are formed from the interview, notes that are made by observing during reading a document, and the like. After, they explained the analytical memos, as those notes that are taken in the process of analyzing and helps the researcher to clearly know about and keep the track of developing the argument or theory as he/she analyzes. It is very applicable in those researches when two or more researchers work on a same data, and also emphasized on using graphs, diagrams, tables, charts that are introduced as visual data displays to organize ideas, producing research report, and identifying patterns.

At the end of this chapter, they bolded the use of computer-aided qualitative analysis that referred to those softwares that do text processing and allow the scholar to segment and code the data. They bolded the computer assisted qualitative data analysis (CAQDAS) networking project that provides information about (CAQAS) packages, such as digital reply system, Atlas.ti, QDA Miner, Max QDA, Nvivo, HyperREASERCH, Transana Qualrus, MiMeG, etc. Digital reply system also does transcribing and concurrently allows that the text and audio files to be displayed, along with supporting the other multimedia formats, such as video and graphics files. These packages also have basic concordancing tools which textual data or phrases and keywords can be explored in a key word in context

(KWIC) format and it is very effective in the situation that the research question relates to the use of language. They stressed on this point that these packages don't do any analysis, and their work is to help to organize and manage data. In other words, in contrast with quantitative packages such as SPSS, that partially do the analysis, they don't perform any real analysis.

The last chapter was about addressing the item of conclusion in research under the title of conclusion: the research report. They foregrounded the items of research, introduction, methods, results, and discussion, which in its abbreviation is IMRAD and as the format for structuring the scientific researches, is very used in empirical studies such as health care and natural sciences. This structure has the advantage of quickly scanning the needed information on the basis of its defined segments. They brought the critique of noble-prize winner and immunologist, Peter Medawar (1996, as cited by Saldanha & O'Brien, 2013, p. 234) as "idealizing that process and misrepresenting the dynamic of thought and discovery that is at the basis of the work reported" (Saldanha & O'Brien, 2013, p. 234), and answered to this criticism by saying that "the research process is much less tidy and linear than the report suggests: the IMRAD structure requires the research to be reorganized so that a 'noise-free' presentation of what are generally seen as the most essential aspects of the study is achieved" (Saldanha & O'Brien, 2013, p. 234). They emphasized on these two parts as the main and important research segments that must be preserved, introduction, and conclusion, while the rest of the reports can be appeared in different forms, but under the condition of the logical sequence must also be preserved which guides the reader in the process of reading the study.

At last, they foregrounded reporting that can be the reporting of quantitative data, which charts, tables, histograms, and barcharts can be useful, and reporting the qualitative data which diagrams, photographs, videos, bullets, and tables can be used for such this reporting. They stressed on asking the question of does it need for participants to remain anonymous in reporting observations, such as when the research uses photographs or videos. They stated that, in small samples the aim is not only reporting percentages, but to give specific values too, and using quantitative method can may lead to the risk of sending the wrong message as there are large differences between the two variables, while the difference in reality may be in 2 or 3 units. If concordances be shown, they must be in KWIC format or key word in co-text and must not be more than 20 or 30 lines as regarding the maximum.

They also bolded using examples for reporting as it is very crucial in product-oriented research for the aim of credibility, making it more readable, and bringing home the points that made. They noted that, the researcher must avoid paraphrasing the example, or simply stating the obvious. They said that, example must be described in detail in such a way that shows itself as the supporter of the claim that made. The example must also be brought with related source of the example for the purpose of credibility. They stated that, in translation field, the degree of difficulty in reporting the outcome of the linguistic analysis highly increases, since two or more different languages are used in this area. They said:

Writers need to make sure readers understand the differences in the two (or more) linguistic systems that are relevant for the texts being discussed without assuming knowledge of any languages other than that in which the report is written. There is no easy way around this challenge, except to describe the relevant aspects of the two systems clearly, providing enough but not too much information so as to avoid confusing readers, and providing good examples with glosses in the language in which the report is written (Saldanha & O'Brien, 2013, p. 242).

References

- Experts think alouds. (n.d.). Retrieved (2019, December 17) from: <https://loft.io/guide/learning/understand/taps/>
- Saldanha, G., & O'Brien, S. (2013). *Research methodologies in translation studies*. London and New York: Routledge.