Motivation, Participation and Organisation in Crowdsourcing Translation – A Case Study on *Yeeyan*

Jun Yang

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

University of Leeds

School of Languages, Cultures & Societies

Centre for Translation Studies

March 2018

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

First and foremost, I would like to give special thanks to my supervisors: Dr Dragoș Ciobanu and Dr Serge Sharoff, for their constant guidance and support. Thanks for guiding me step by step throughout this journey, for putting this project into shape, and for showing me the best qualities as a researcher. I consider myself to be extremely lucky to have had two incredible supervisors.

I am grateful to Dr Alina Secară for her suggestions to my research, and for her friendship.

I would also like to express gratitude to my examiners – Professor Andrew Rothwell and Dr James Wilson – who have been so inspiring in guiding the directions of my future research and for their suggestions that have helped improve this thesis.

I am very grateful to the *Yeeyan* community. Thanks to Dr Jiamin Zhao, the editors – Yuling Wang, Xin Xu, Jingjing Guo and Xiaofei Chen, and the technician groups for offering kind help during my data collection. I would like to thank all the *Yeeyan* members for creating such a good resource for research. And special thanks to my participants who were so willingly to take part in this project.

Many thanks to Raisa McNab and Kathy Walters in *Sandberg Translation Partners Ltd* for kindly offering me the project management training and for providing me the opportunity to learn real-life experience in the Language Services Industry. The training outcome shed light on the comparison between professional industry and the crowdsourced environment in my study.

I would also like to thank Dr Vivian Ran Xu for her expertise on terminology preparation; to David Yu Yuan for sharing his knowledge in translation quality assessment and English-Chinese error typology.

A special ‘thank you’ to my dearest friend Xi Wang for accompanying me throughout this long journey.

Last, but by no means least, my gratitude goes to my family for their unconditional love and encouragement, and for supporting me to achieve my goal.

Abstract

Crowdsourcing, i.e. engaging online communities in productive activities, has become a commonly used practice in content-generating areas, including translation. It has fostered alternative translation models, such as volunteer or paid crowdsourcing by professional or bilinguals translating different text types with various purposes. Scholars embarked on the investigations of crowdsourcing translation practices in the West (Dolmaya, 2012; Jiménez-Crespo, 2013a; O'Brien and Schäler, 2010; Olohan, 2014, 2012; Orrego-Carmona, 2015). However, little focus has been paid to crowdsourcing translation in China. This study explores the mechanism of crowdsourcing translation drawing from one of the largest Chinese amateurs' communities – *Yeeyan,* seeking to raise awareness of the particular collaborative model both in the Language Services Industry and in Translation Studies.

The study examines collaborative translation projects in *Yeeyan*, aiming to find out:

1. the motivating and de-motivating factors that influence participation;
2. the individuals' contribution to crowdsourcing translation projects;
3. the quality of crowdsourcing outputs and the effectiveness of quality control methods;
4. the texts types that are feasible to translate via crowdsourcing;
5. the peer-learning process taking place in the crowdsourced environment.

Questionnaire survey and *observation* were the main methods for data collection.

The result of the *questionnaire survey* suggested that individuals are usually driven by multiple reasons both intrinsically and extrinsically. However, extrinsic motivations are perceived to be more important than intrinsic motivations. Moreover, it revealed that the individual's behaviour (i.e. the time and effort that one invests in crowdsourcing translation) is influenced by these motivations.

Empirical data of collaborative translation projects in the chosen environment

was collected through *Observation*. *Workflow Mapping* demonstrated how projects progress, and identified two types of workflow that cater for different translation purposes:

- the user-driven flexible workflow mainly depending on the participants' spontaneity;
- the waterfall workflow strictly managed and monitored by the project manager.

Furthermore, the study highlighted the role of communication in facilitating collaboration. *Dialogue Act Analysis* was conducted to find out the frequency and category of interaction, and the learning process embedded in online peer-to-peer interaction. *Social Network Analysis* was then used to visualise the communication patterns, to characterise the power relationship between participants, and to illustrate the individuals' contribution to the projects.

An *Error-based Quality Assessment* dealt with the outputs of crowdsourcing translation at different stages (translation, revision, finalisation). The result revealed that the crowd can produce 'fit-for-purpose' translation through collaboration. It proved that it is feasible to translate both specialised texts and literary texts through crowdsourcing. The assessment focused on the effectiveness of quality control methods and showed that 'peer-revision' is the most efficient approach in improving translation quality. In addition, the study identified the significant error types which the crowd makes most often. More importantly, a peer-learning process was recognised in both the interaction between participants and the feedback provided during collaboration.

The study produces a comprehensive evaluation of collaborative translation projects in the crowdsourced environment. The outcome is expected to bridge the gap between this relatively new phenomenon in translation practices and Translation Studies, to offer insights into how crowdsourcing translation projects can be organised more effectively, and to offer suggestions regarding setting up enhanced crowdsourcing initiatives for both translation production and translation training.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BP I	Business Project I
BP II	Business Project II
CAT	Computer Assisted Translation
DAA	Dialogue Act Analysis
FOSS	Free Open Source Software
GP	Gutenberg Project
HP I	Health Project I
HP II	Health Project II
LSP	Language Service Provider
MT	Machine Translation
PM	Project Manager
SNA	Social Network Analysis
SL	Source Language
ST	Source Text
TEP	Translation – Editing – Proofreading
TL	Target Language
TT	Target Text
UGC	User-generated Content
Abbreviations in the Interaction Typology	
AC	Availability clarify
AK	Acknowledgement
APM	Apologise for mistakes/problems
AR	Availability request
ATA	Assistance arrangement
ATR	Assistance request
CF	Confirm the receipt of file
CM	Commission
CT	Comment
DC	Deadline clarify
DLR	Delay report
DR	Deadline request
DVR	Deliverables request
FC	Financial-related clarify
FFC	File format clarify
FFR	File format request
FQ	Financial-related request
GIF	General information
IC	Information clarify
IFM	Inform financial mistakes/problems
IMD	Inform mistakes/problems in deliverables

IR	Information request
IRM	Inform reference materials
IV	Invoice
MT	Motivation
PI	Project instruction
PIC	Project instruction clarify
PIR	Project instruction request
PO	Purchase order
PR	Problem report
QT	Quote
QTN	Quote negotiation
RA	Role accept
RC	Revision comment
RD	Role decline
RDS	PM reminder of deliverables submission
RH	Role hold
RQ	Role request
RR	Revision request
SC	Status check
SO	Social interaction
SR	Status report
STR	Solution to problem
TA	Task assignment
TAM	Term agreement
TC	Term clarify
TD	Term discussion
TR	Term request
TS	Task submission
XAM	Text agreement
XC	Text clarify
XD	Text discussion
XR	Text request
Abbreviations in the Error Typology	
AD	Addition of information
AM	Ambiguous translation
AW	Awkward syntax
CS	Character spelling
DI	Meaning distortion
LT	Literal translation
MI	Miscellaneous
NN	Not idiomatic, non-native use of target language
OM	Omission of information

PU	Incorrect punctuation
RE	Problems with register
TE	More suitable term available
TI	Inconsistency within the text
WTE	Term not correct

Chapter 1 Introduction

1.1 Background and research questions

The development of globalisation has promoted the rapid growth of people's demand for cross-cultural communication. Technological connectivity has made information exchange easier and more frequent. In the past, when information and communication technology was underdeveloped, the introduction of foreign cultural content mainly relied on the official channels, i.e. television broadcasts, cinemas, DVD rentals and purchases and publishing houses. It took complicated procedures and a long production cycle until the content would become accessible to audiences. However, these outdated forms of cross-cultural transmission no longer meet the audiences' needs today. Many prefer the global trend of 'synchronous communication', in which the problem of time-lag has been largely solved. For example, fan subtitling groups would localise popular cultural products (e.g. American, Korean and Japanese dramas) within hours or days from the live broadcast.

The current internet-mediated communication has offered fast and easy ways of content consumption. It broadens the public's horizons and reinforces their curiosity and the needs towards foreign cultures, especially in the Chinese group when they have experienced long-term information control. The growing trend is that people are no longer satisfied with the content of domestic information, as well as the traditional forms of communication. However, because not everyone has sufficient knowledge of foreign languages, the first obstacle for cross-cultural communication is the *language barrier*, followed then by the eagerness to receive *timely information*.

Collective intelligence solved the two problems by involving different social actors in the communication process. **Crowdsourcing**, i.e. engaging online communities in productive activities, came into being. It is a feasible and suitable solution as it exploits a huge reservoir of skills and competences. The once professional-dominated business has become an inclusive field that allows amateurs, trainees, fans, and activists to work on the content

production and introduction process. On the one hand, it is the consequence of the imbalanced demand-supply relationship. The explosion of content to be processed surpasses the available resources in the professional industry. On the other hand, people are empowered with the opportunity to decide what they wish to consume and to produce fit-for-purpose content.

In the context of Chinese society, the growing number of students in higher education has resulted in a crowd with foreign language competence, who are capable of performing translation activities to overcome the language barrier in cross-cultural communication. More importantly, traditional Chinese education stresses continuously upon students the value of ‘serving the public good’. The idea of ‘sharing’ is deeply-rooted and therefore, in addition to their main occupation, these educated people seek opportunities that satisfy their personal interest while doing something ‘meaningful’ for the society.

In 2017, there were 252 Chinese universities offering translation as the core course in undergraduate education, and 215 universities offering Master programmes in translation and interpreting studies (China National Committee for BTI Education, 2017). This reality explains another characteristic of crowdsourcing translation in China: the huge number of translation/interpreting students need crowdsourcing platforms to practise their developing skills. CATTI (*China Accreditation Test for Translators and Interpreters*) is a national standard test that provides qualification certificates of translation and interpreting. It is an important test that employers in the Language Services Industry value, especially for state-owned institutions. According to an official report, from 2003 (the year when the test was implemented) to 2017, the total number of applicants surpassed 720,000, of which only 12% – 87,000 applicants – gained the qualification certificates (National Translation Test and Appraisal Centre, 2017).

If translation students are determined to enter the Language Services Industry, it is clear that the training they received at university is not sufficient. Crowdsourcing provides the opportunity for students to gain real-life translation experience and prepare for the challenges that they may

encounter in their future career. Hence, there are large number of trainees who are keen on contributing to crowdsourcing translation, which supplies the important labour force to the crowdsourcing initiatives.

As mentioned earlier, fan subtitling was the first activity that motivated online communities to collaborate on translation tasks. It fulfils Chinese consumers' needs for cross-cultural communication from the perspective of entertainment. However, the need for cultural transmission does not stop at entertainment. Mass foreign media content, and other various forms of foreign information are in acute demand, which leads to a variety of crowdsourcing translation initiatives.

In general, there are two types of crowdsourcing initiatives in China:

- Crowdsourcing portals[1], e.g. *Taskcn*, *Zhubajie*, *Zuodao*, and *Trycan* They work in the same mechanism as *Amazon Mechanical Turk*, in which the crowdsourcer distributes tasks, which the crowd subsequently perform for monetary rewards. Low cost and high speed are usually expected by the crowdsourcer.
- Information-oriented crowdsourcing communities[2], e.g. *CNPolitics*, *Guokr*, *Hupu*, *Fiberead*, and *Yeeyan*. These communities usually have their own focus. *CNPolitics* is an independent volunteer group committed to introducing academic studies on Chinese politics to the Chinese public. *Guokr* is a scientific-oriented community: volunteers contribute to the dissemination of science and technology knowledge to the general public. *Hupu* is a sports community in which volunteers and fans gather to translate NBA news and share other sports-related content to the Chinese audiences. *Fiberead* is the newest

[1]Crowdsourcing portals: Taskcn is available from http://www.taskcn.com/; Zhubajie is available from http://www.zbj.com/; Zuodao is available from https://www.zuodao.com/; Trycan is available from http://www.trycan.com/fy_public.html; Amazon Mechanical Turk is available from https://www.mturk.com/.
Taskcn and Zhubajie are two general crowdsourcing portals that cover all kinds of services, including translation. Zuodao and Trycan are translation-specific crowdsourcing portals.

[2]Information-oriented crowdsourcing communities: CNPolitics is available from http://cnpolitics.org/; Guokr is available from https://www.guokr.com/; Hupu is available from https://nba.hupu.com/; Fiberead is available from https://fiberead.com/; Yeeyan is available from http://g.yeeyan.org/.

> crowdsourcing community that focuses on book translation and digital publication. It claims to help authors reach global market places without heavy investment. Foreign authors who wish to have their book published in Chinese make use of the community. Translation amateurs and language learners are the main labour force. The highlight of this community is that translators and authors are able to interact with each other to ensure the quality of translation. Preselection of participants is needed in the community, and authors and translators receive royalties. Last, but by no means least, *Yeeyan* is the earliest crowdsourcing translation community in China. With the aim of introducing 'foreign essence' to China, the translations completed on this platform cover a wide range of topics, including business, health, technology, media, and culture. The community is very active in book translation and digital publication, as well.

Crowdsourcing translation is becoming more and more popular in China, as it has changed the traditional direction of information flow. In contrast to the top-down communication model, in which audiences are passive receivers of what is released by the authority, crowdsourcing translation ensures bidirectional communication, in which the audiences are both information producers and receivers. For **information-oriented crowdsourcing initiatives**, it provides an autonomous environment in which the general public decides the content they wish to be imported from the foreign world. Moreover, unlike translation services in the professional industry which sometimes tailor the output for specific target groups, information-oriented crowdsourcing translation reaches a much wider range of audiences. Therefore, in terms of cross-cultural communication, it enhances the social influence of translation.

Scholars have already embarked on the investigation of crowdsourcing translation in the West. For instance, the motivations for participating in crowdsourcing translation (Dolmaya, 2012; O'Brien and Schäler, 2010; Olohan, 2012, 2014), and the acceptability and adequacy of crowdsourcing translation outputs (Deriemaeker, 2014; Jiménez-Crespo, 2013a; Orrego-Carmona, 2015) have been investigated recently. However, crowdsourcing

translation practices in China have not received enough attention from Translation Studies scholars to date.

Chesterman (2014:82) mentioned a view held by relativists that the field of contemporary Translation Studies is 'too Eurocentric or too Western', and it needs to expand and incorporate non-Western approaches in order to address more adequately 'some current trends that challenge the development of translation theory'. The relativists argue that the current Eurocentric views are not suitable for some translation practices in non-Western cultures, for instance, the increase of non-professional and group translation practices, crowdsourcing and fan translation in particular. The issue is caused by the different perceptions towards translation: some see Translation Studies as an empirical human science which needs standardisation, and some see it as a hermeneutic discipline in which conceptual argument about meaning and interpretation is more valued. Chesterman (ibid:88-89) emphasised that '[a]s translation is a cultural phenomenon, we can expect a huge variety of views and approaches. […] What matters is whether a concept, model, hypothesis or theory turns out to be useful or not, in solving a given problem and/or in bringing deeper understanding. Or, indeed in generating further fruitful questions.'

I share Chesterman's view and believe that there is no superior idea just because it has its origins in a particular culture or region. However, as mentioned earlier, crowdsourcing translation in the Chinese context has its specificities. In addition, considering that China is one of the largest markets for content consumption and has a huge demand for translation services, an investigation of the phenomenon is of special importance.

In order to address these specificities and also widen the field of Translation Studies more generally by including one of the necessary Chinese perspectives, this study explores crowdsourcing translation by analysing empirical data obtained from the Chinese amateur community *Yeeyan*. There are two reasons why *Yeeyan* was chosen. First of all, unlike other information-oriented crowdsourcing communities, which only focus on one particular type of content, *Yeeyan* is a much more diverse community that

encompasses a variety of source text categories. Secondly, *Yeeyan* brings together an open portal and an open community, meaning that most of the translations produced in *Yeeyan* are publicly accessible, and thus ideal for data collection and research.

Founded in 2006, *Yeeyan* is the largest online translation community and crowdsourcing translation platform in China, with more than 600,000 registered users. Up to now, over 410,000 articles and around 300 books have been translated and e-published here. Two major types of translation are performed in *Yeeyan*: article translation carried out for the purpose of knowledge sharing, and book translation that is performed for publication. In *Yeeyan,* a source text (ST) repository is compiled by the employed editors, who also ensure the legitimacy of any translation performed and published through the *Yeeyan* community. The ST repository is accessible to all registered users who can choose to translate either independently or to collaborate with others. In the individual independent translation mode, the process relies on self-revision for quality assurance, as well as the post-feedback provided by the audiences. Collaboration is a common method when translating large texts in which self-revision and peer-revision are implemented for quality control. Here, the definition of ‘collaboration’ is in accord with its conventional meaning: each individual performs a small fraction of the activity. To be more specific, in a collaborative project, each participant translates a part of the ST; peers can view, comment and revise others’ translations. *Gutenberg Project* is the signature product of the *Yeeyan* community, which endeavours to translate and publish books that are already available in the public domain. Preselection of participants and strict monitoring are needed in the *Gutenberg Project* to meet the quality criteria for publication. Moreover, translators can either work individually or collaboratively, depending on the personal willingness and the number of qualified participants.

Since the strength of crowdsourcing relies on the collective intelligence (Howe, 2006), this study focuses on collaborative translation projects in the *Yeeyan* community, attempting to find out whether the claimed ‘crowd wisdom’ can produce quality translation. To gain a comprehensive picture of

how crowdsourcing translation works, the study examines both types of translation (article translation done for knowledge sharing and book translation done for publication), as they represent different settings in text genre, purpose of translation, the degree of organisation, and project management.

Anastasiou and Gupta (2011:641-642) pointed out the major challenges for crowdsourcing translation include quality, motivation, and control. *Quality* is influenced by the competence of the participants as it contains the feature of 'amateurship', and the community which translates does not always consist of accredited, certified and experienced translators. In most of the cases, it gives equal opportunities to professional translators, less-experienced and non-professional people to be involved in linguistic activities. *Motivation* is what makes crowdsourcing feasible. If the motivation is missing, it will be difficult to gain the crowd's attention and make them emotionally attached to a desired task. *Control* refers to the coordination and management of the community; it can be a lead user with high motivation or some in-house community managers who can take on the role of moderating the collaboration between the crowd.

Inspired by Anastasiou and Gupta's (2011) work, this study gathered and analysed empirical evidence in order to provide a comprehensive evaluation of the approaches taken within the *Yeeyan* community to tackle the claimed obstacles of 'quality, motivation and control'. Moreover, because of the specificities of the Chinese context, the chance of sharing foreign information is what makes crowdsourcing translation attractive. The openness of *Yeeyan* encourages all kinds of contributions which support the full-functioning of the community. It is necessary to find out how exactly individuals' contributions influence the translation product, as well as the community. Furthermore, the potential of using crowdsourcing initiatives to practise translation skills motivates many trainee translators to participate. It is also worth investigating whether a learning process is embedded in crowdsourcing translation.

To sum up, this study focuses on the following research questions which are investigated in the specific context of the Chinese *Yeeyan* community:

- Why do people participate in crowdsourcing translation projects?
- How is the individual's contribution valued?
- What is the quality level of crowdsourcing translation, and how is quality controlled?
- What kind of texts can be feasibly translated through crowdsourcing?
- Does participating in crowdsourcing translation have a clearly-defined learning dimension? If so, what are its features?

The outcome of the study is expected to bridge the gap between Translation Studies and the actual practices of the new translation phenomenon – crowdsourcing translation in the Chinese context. It also aims to suggest possible approaches to designing and organising effective crowdsourcing translation projects, and to offer insights into how crowdsourcing initiatives can be put to better use in translation production, as well as translation training.

1.2 Thesis structure

Chapter 1 of the thesis serves as an introduction. It includes contextual background of the study, the researcher's motivation, as well as the research questions.

Chapter 2 provides a theoretical framework of the study. It reviews literature related to the fundamental issues in the project. This chapter starts with the clarification of key concepts, including crowdsourcing and crowdsourcing translation. It explains why crowdsourcing translation is possible from the perspective of human ability, technical support and social culture. It then summarises motivation studies in crowdsourcing translation, which provides a basis for assumptions about the first research question. The chapter also gives a broad overview of project management in Translation Studies with emphasis on workflow and communication management, which is used as a reference when examining the participation and organisation of crowdsourcing translation projects. Finally, it discusses the concept of translation quality and presents the approaches to translation quality evaluation in both academia and the professional industry, which generates

the methodological and theoretical foundation of the translation quality assessment in this study.

Chapter 3 introduces the main methodological approaches of this study. It explains the metrics and considerations in the experiment design. Each method (*Questionnaire survey* and *Observation*) is described individually to specify its suitability for the study and the type of data it enabled the researcher to collect. Three-step procedures of data analysis in accordance with the types of collected data (communication record and translation texts with edit history) are explained.

Chapters 4-7 analyse and discuss the collected data on the *Yeeyan* community. **Chapter 4** presents the findings of the *questionnaire survey*. It first reports on the demographic characteristics of the *Yeeyan* community and answers 'WHO' is contributing to crowdsourcing translation. Motivations and other influential factors are presented when discussing the 'WHY' question in the theoretical framework. Findings of the self-declared participation are mentioned when analysing the 'HOW' aspect of crowdsourcing translation.

Chapter 5 reports the findings of *observation*, which elaborates on the actual practices of crowdsourcing translation projects. The result is compared with the self-declared findings from the questionnaire survey. It first maps the workflow of the observed project, then assesses the participants' productivity in performing translation-related tasks. Based on the participants' activities in the observed projects, this chapter summarises characteristics in the workflow patterns and organisational structures. The learning perspective of crowdsourcing translation is discussed towards the end of the chapter.

Chapter 6 analyses the communication data collected from *observation*, as peer-to-peer interaction is recognised as an important feature in collaborative translation projects in the crowdsourced environment. It explores the frequency and category of interactions through *Dialogue Act Analysis*. The chapter then investigates the relationships and information flow between the participants through *Social Network Analysis*. In order to establish comprehensive communication patterns and criteria that enhance the

collaboration in translation projects, the findings of communication analysis are compared with a *Students' Team Project*, which replicates multilingual localisation projects in the professional industry. The chapter evaluates the participants' performance and the collective efforts reflected in peer-to-peer interactions.

Chapter 7 explores the translation outputs of the chosen crowdsourcing projects through *error-based quality assessment*, and it also examines the effectiveness of the quality control methods implemented. Moreover, the chapter explores the participants' competence through the quality evaluation of the three versions of translation outputs (draft translation, self-revised version, and peer-revised version).

Chapter 8 summarises the key findings and implications of this research, and it also offers suggestions for future organisation of crowdsourcing translation. The chapter then discusses the contributions of the study and identifies its limitations. Suggestions for future research bring the thesis to a close.

A list of references and appendices follow the conclusion.

Chapter 2 Literature Review

2.1 Crowdsourcing in general

2.1.1 The definition of crowdsourcing

The term ‘crowdsourcing’ was coined by Howe (2006a) and it was defined as ‘the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. [...] The crucial prerequisite is the use of the open call format and the large network of potential labourers.’ This definition highlights that the ‘crowdsourcing’ process is both initiated and sponsored by an organisation, and it involves a large number of people – the ‘crowd’ – which is directed or managed by the organisation throughout the process, too.

Many scholars have since attempted to propose their own definitions of crowdsourcing based on a diverse set of practices and a number of different theoretical considerations, and some of them even have several versions of such a definition: for instance, Howe (2006a, 2006b, 2008) has three versions, while Brabham (2008a, 2008b) has two versions. Zhao and Zhu (2012:421) conducted an evaluation of crowdsourcing research by surveying the landscape of existing studies and pointed out that the definitions given by these scholars tend to focus on ‘some perspective[s] or feature[s] of crowdsourcing in a particular area, and it is difficult to reach a consensus’. For instance, Brabham (2008b:79) put forward a definition as ‘a strategic model to attract an interested, motivated crowd of individuals capable of providing solutions superior in quality and quantity to those that even traditional forms of business can’ when he tries to introduce crowdsourcing through some case-studies and applications that emphasise the potential to exploit the crowd of ‘innovators’ for ‘problem-solving’. In his later research, Brabham (2013) enriched the definition by adding two more ingredients – first of all, an online environment that allows the work to take place and the crowd to interact with the organisation; and secondly, the mutual benefit for the organisation and the community.

To get an exhaustive and consistent definition, Estellés-Arolas and González-Ladrón-de-Guevara (2012) first analysed the existing definitions in order to extract the common features, then created an integrated definition, and finally tested it on a wider range of crowdsourcing applications. The outcome is as follows:

> *'Crowdsourcing is a type of participative online activity in which an individual, an institution, a non-profit organisation, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcer will obtain and utilise to their advantage what the user has brought to the venture, whose form will depend on the type of activity undertaken.'* (Estellés-Arolas and González-Ladrón-de-Guevara, 2012:197)

In this PhD project, I build on Estellés-Arolas and González-Ladrón-de-Guevara's definition as it covers the majority of existing crowdsourcing practices. Their analysis of 32 definitions proposed by scholars in various domains highlighted eight characteristics common to any given crowdsourcing initiative:

> *'the crowd; the task at hand; the recompense obtained; the crowdsourcer or initiator of the crowdsourcing activity; what is obtained by them following the crowdsourcing process; the type of process; the call to participate; and the medium'* (Estellés-Arolas and González-Ladrón-de-Guevara, 2012:198).

When observing and analysing translation projects in the *Yeeyan* community, special attention was paid to the aforementioned characteristics in order to provide a comprehensive evaluation of the *Yeeyan*-specific crowdsourcing

mode and highlight the aspects in which it differs from other implementations of crowdsourcing platforms.

Despite there being significant differences between ‘crowdsourcing’ and ‘outsourcing’, the two terms are still frequently confused. *Outsourcing* refers to ‘the use of external agents to perform one or more organisational activities, reflecting a company contracting other companies to provide services that might otherwise be performed by in-house employees’ (Zhao and Zhu, 2012:421). Of great importance in this context is the presence of the term ‘contract’ (ibid). In *outsourcing*, the client company looks for a supplier and sets the requirements, then the preselected supplier provides the client company with products/services, in accordance with the contract. However, in *crowdsourcing*, a ‘contract’ between the labour and the company is rarely included. The crowdsourcer (e.g. client company) issues an open call and individuals within the crowd (e.g. an online community) provide inputs to the crowdsourcer on a voluntary basis. It is difficult (but not impossible), to secure a contract because that would entail ratifying an accord with multiple parties who are often anonymous.

Another difference rests with the relationship between the labour and the client company. *Outsourcing* largely depends on business relationships (i.e. financial incentives), while *crowdsourcing* may have a much more diverse and friendlier user-client relationship due to the participation motivation, which may lead to multiple incentives (ibid). Having said that, one should acknowledge also that, due to the additional complexities which financial recompense brings to the crowdsourcing process, intermediaries have appeared in the online medium in order to handle this still significant crowd motivator – e.g. crowdsourcing portals that take payments from the clients and distribute them to the crowd participants.

2.1.2 Crowdsourcing framework and relevant research

2.1.2.1 Theoretical framework on crowdsourcing

Scholars have shown interest in crowdsourcing in multiple disciplines, including:

- communication studies – e.g. the emergence of the citizen journalism and the generation of online-media content (Aitamurto, 2016; Domingo et al., 2008; Muthukumaraswamy, 2010);
- social-psychology studies – e.g. motivations in participating crowdsourcing activities (Borst, 2010; Brabham, 2008a, 2012; Dombek, 2014; Hars and Ou, 2002);
- business studies – e.g. using crowdsourcing to integrate customer/user for open innovation that addresses both value creation and customer-firm interaction (Chesbrough, 2003; Leimeister et al., 2009; Trompette et al., 2008); and
- computer science studies – e.g. crowdsourcing for the development of open-source software and the design of a crowdsourcing system that promotes participation (Doan et al., 2011; K. Lakhani and Wolf, 2005; Leimeister et al., 2009) – to name but a few domains involving the use of crowdsourcing.

In Zhao and Zhu's (2012) evaluation of crowdsourcing research, they found that there is a lack of a theoretical foundation for the analysis of this type of activity, which indicates the immaturity of this research area: out of 55 academic papers in their collection pool, only 10 used non-empirical methods that focus on the conceptualisation of crowdsourcing; on the other hand, 16 studies used empirical data gathered through lab and field experiments (10 and 6, respectively). The rest of them used interpretative, survey and secondary data to explore crowdsourcing events. Such diversity in crowdsourcing research can be explained by the fact that crowdsourcing is still a new research area and many scholars focus on its applications in various contexts rather than its conceptual foundation. Interestingly, Zhao and Zhu (2012) paid little attention to crowdsourcing in the translation

domain, and only three studies (Ambati et al., 2010; Callison-Burch, 2009; Munro et al., 2010) were mentioned in the area of crowdsourcing for natural language processing and machine learning.

Zhao and Zhu (2012) presented a conceptualisation framework of crowdsourcing by identifying the fundamental dimensions and their relationships based on Malone et al.'s (2010) work on harnessing the crowd for collective intelligence. As displayed in Figure 1, the framework addresses a set of key questions in crowdsourcing:

1. Who is performing the task?
2. Why are they doing it?
3. How is the task performed?
4. What about the ownership and what is being accomplished?

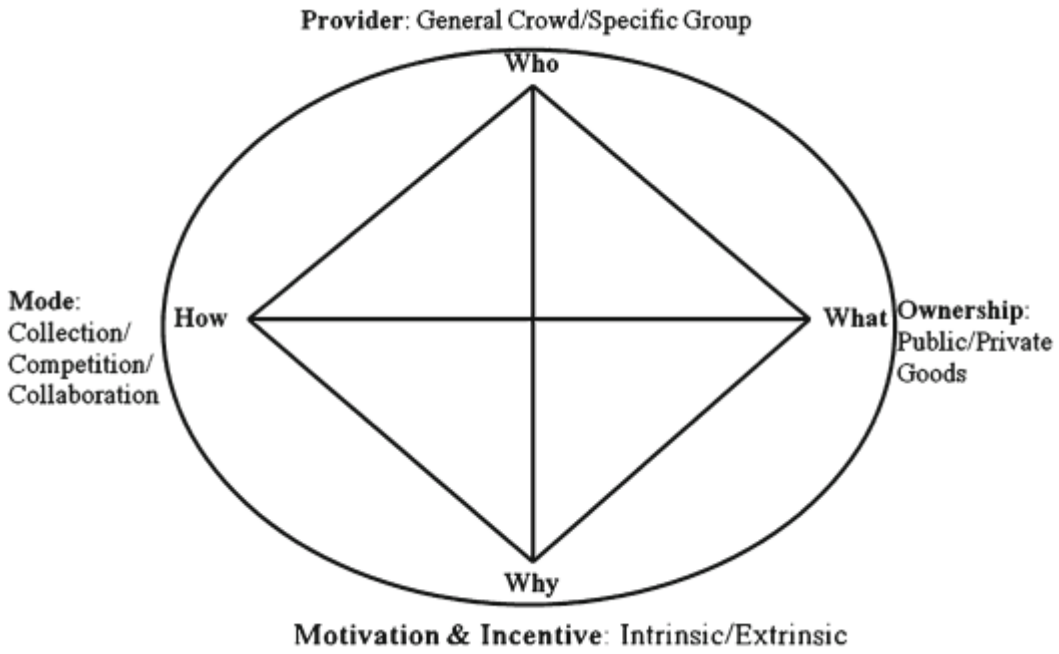

Figure 1. Theoretical framework of crowdsourcing: fundamental dimensions. Zhao and Zhu (2012:422) adapted from Malone et al., (2010).

The four dimensions overlap with some of the characteristics of crowdsourcing summarised by Estellés-Arolas and González-Ladrón-de-Guevara (2012) (see section 2.1.1), except that they paid less attention to the ownership of task outcome – i.e. considering the outcome of crowdsourcing

task as goods, and classifying them as either public goods or private goods (Kazman and Chen, 2009).

The first dimension – 'Who' – explains the attributes of the crowd: people who perform the crowdsourcing task are either undefined general crowds or specific groups.

The second dimension – 'Why' – addresses the motivations of participation: extrinsic reasons and intrinsic reasons which are discussed in detail in section 2.3.

The third dimension – 'How' – addresses three categories of structures/processes of the task: *collection* – activities can be divided into small pieces that can be done independently of each other; *competition* – only one or a few good solutions are needed and chosen from the crowd; *collaboration* – each individual performs a small fraction of the activity. My study investigates the 'collaboration' category in the process of crowdsourcing translation which is discussed in Chapter 5.

The final dimension – 'What' – looks at the attribute of the task outcome, discussing if it can be considered either a public or a private product.

Rouse (2010) put forward a preliminary taxonomy of crowdsourcing, which focuses on the different capability levels of crowdsourcing suppliers, different motivations, and different allocation of benefits. These three dimensions can be well mapped to the Zhao and Zhu's (2012) theoretical framework, i.e. 'What' – supplier capabilities/nature of the task; 'Why' – forms of motivations. The third dimension in Rouse's taxonomy – allocation of benefits – corresponds to one of the characteristic proposed by Estellés-Arolas and González-Ladrón-de-Guevara (2012) – mutual benefits – and it also overlaps with the 'Who', 'Why', and 'How' dimensions in Zhao and Zhu's (2010) model.

This research uses the above-mentioned theoretical framework and attempts to find out the features of the 'Who', 'Why' and 'How' dimensions in *Yeeyan*, as well as propose a separate, novel dimension – 'Quality' – of the crowdsourcing output.

2.1.2.2 Research on crowdsourcing

Zhao and Zhu (2012) classified crowdsourcing researches from the following three different perspectives: the participant's perspective, the organisation's perspective, and the system's perspective (Figure 2).

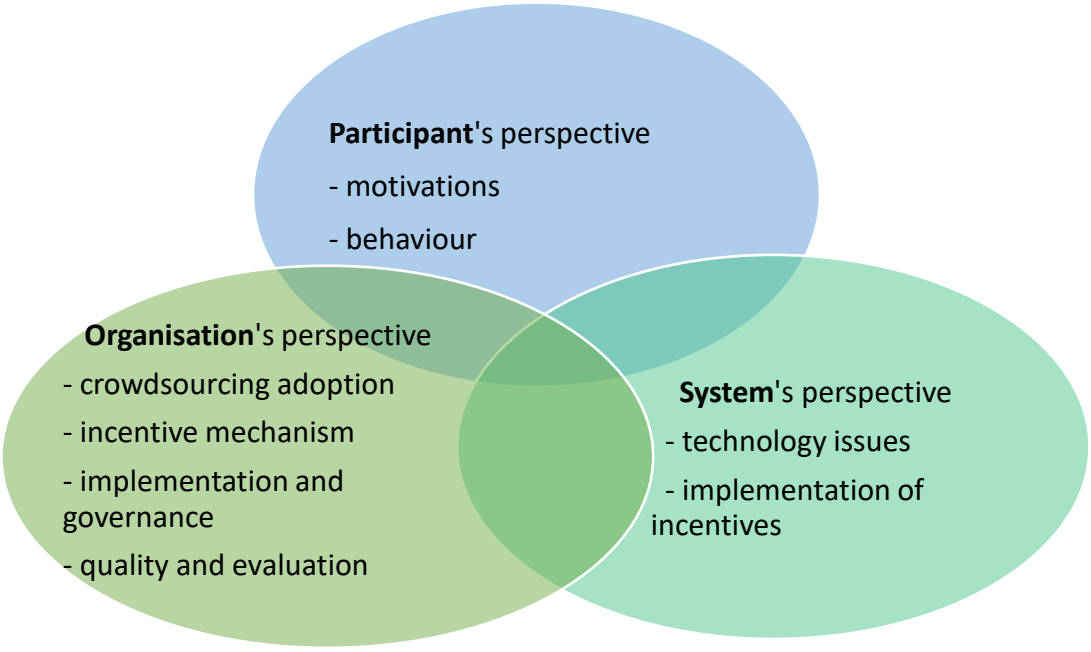

Figure 2. Directions for crowdsourcing research – adapted from Zhao and Zhu (2012:426)

The participant's perspective

The participant's perspective includes social-psychological studies that explore people's perception and motivations for participation in crowdsourcing activities, which facilitates the sustainable development of crowdsourcing communities. Motivations in different contexts may differ. For example, people may have different reasons for contributing to a *business-oriented* project compared to a *public-goods* project (Brabham, 2008a). Understanding why people are willing to participate, as well as their influencing factors, can shed light on the organisation of crowdsourcing activities and the design of crowdsourcing systems. The motivations to participate are discussed in detail in section 2.3. The participant's behaviour

is believed to be closely related to the crowd's quantity and quality of contribution (Zhao and Zhu, 2012).

Regarding the issue on the crowd's effort and *quantity* of contribution, Stewart et al. (2010) proposed a SCOUT model to reflect the participation inequality. Through an analysis that measures the quantity of contributions correlated with responses to motivation and incentives, they formalised three components of the SCOUT model:

- (S)uper contributors – those who consistently give super effort in terms of quantity and are driven mainly by altruism (intrinsic reward);
- (C)ontributors – those who provide moderate effort in terms of quantity and are mainly driven by extrinsic rewards, and
- OUT(liers) – those who only provide low-level effort that are not sufficient for receiving an award.

My research attempts to correlate the motivations with the efforts put into crowdsourcing translation-related activities using the self-declared data collected from the questionnaire survey. The result is in contrast with Stewart et al.'s (2010) model in that people with extrinsic motivations are more likely to invest more time in translation-related activities in the crowdsourced environment (details see section 4.3 in Chapter 4).

Rajala et al. (2013) explored the diversity of the participating users, and highlighted the importance of engaging *lead* users in crowdsourcing initiatives for several reasons:

- first of all, lead users are passionate about collaborating with the company and are most willing to participate because there is potential value created for their own needs in the participating process;
- secondly, they are highly capable of using the products and have a good insight into the products.

The influence of lead users is reflected in this study, as well (for details, please see 6.2.3 in Chapter 6). Overall, the above studies highlighted the crowd's diverse effort of contribution to crowdsourcing activities with regard to different types of participants.

The organisation's perspective

Geiger et al. (2011) developed a meta dimension for their crowdsourcing taxonomy which focuses on the mechanisms that impact the process and can be controlled by the crowdsourcer. As can be seen from Figure 3, the four dimensions are also covered by Estellés-Arolas and González-Ladrón-de-Guevara's (2012) definition discussed in section 2.1.1.

- The first dimension, 'preselection of contributors' decides **who are** the participants and the type of **the call** for participation.
- Secondly, 'remuneration of contribution' stresses the tangible **benefit** gained by the crowd.
- Thirdly and fourthly, 'aggregation of contributions' and 'accessibility of peer contributions' are factors that influence the type of **process**.

All of these dimensions are included in the organisation of crowdsourcing.

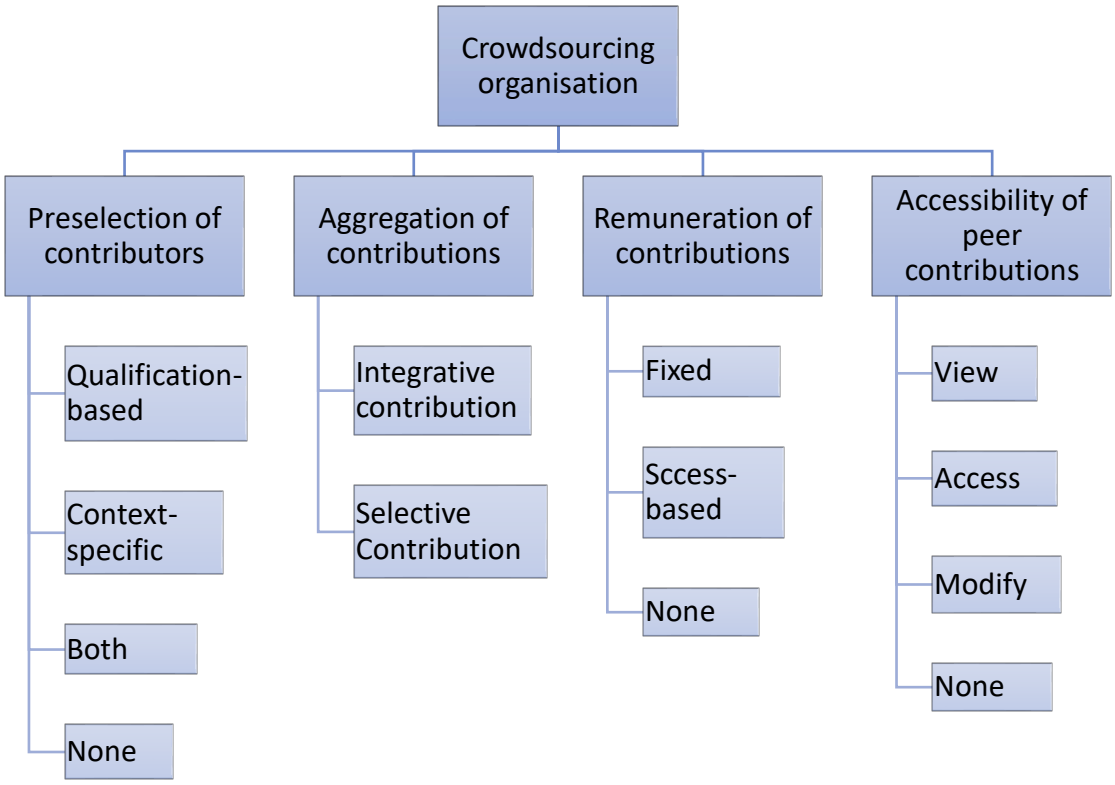

Figure 3. Crowdsourcing organisation – summarised from doctoral thesis (Morera-Mesa, 2014:38-39).

Contributors to the crowdsourcing task can be either preselected or not, and there are usually two approaches in the preselection process: ‘qualification-based’, meaning that participants need to prove their capability in advance – for example, by taking a sample test; or ‘context-based’, meaning that participants need to meet certain context conditions – for example, being a user of a product/service.

In the context of crowdsourcing translation, ‘preselection of contributors’ is often applied when ‘good’ quality is required (e.g. in *Yeeyan*’s *Gutenberg Project*, the participants are expected to have the competence to produce translations that meet the criteria for publication) or certain attributes are desired (e.g. in *Free Open Source Software* community, knowledge about the software is usually expected). My research demonstrates that, while ‘preselected’ contributors do have higher competence than the general translators, their outputs still need to be vigorously revised and the quality assurance stage has to rely on other procedures (e.g. peer-revision) (see 6.4).

Two types of contribution have been investigated so far in the literature: ‘integrative contribution’ and ‘selective contribution’ depending on how the contribution is valued: collectively or individually. Integrative contribution means that the complementary input from the crowd is pooled; selective contribution represents a comparison between individual contributions with the aim of selecting some of them.

The subject of this study – *Yeeyan* – features integrative contribution both in the operation of the community and in the process of crowdsourcing translation (discussed in section 4.5.1 in Chapter 4 and Chapter 5). Selective contribution cannot work in information-oriented crowdsourcing initiatives like *Yeeyan* due to the large amount of source content to be processed. More importantly, it is against the nature of community building if only particular contributions are valued (discussed in section 2.2.2.2).

The types of ‘*remuneration for contributions*’ are classified according to how the contributors’ inputs are treated. Fixed remuneration indicates that participants are treated homogeneously, while success-based remuneration

indicates that the participants are treated heterogeneously, and only the selected contributions will receive a payment. There are also cases where no remuneration is given, but it only means the absence of monetary benefits; participants can gain other benefits according to their motivations for participation. Commercial crowdsourcing portals usually apply fixed remuneration for contributors. In the *Yeeyan* community, two categories of remuneration were found: no remuneration for general translation projects that aim for knowledge-sharing; and success-based remuneration for translation projects that are done for publication in which contributors gain royalties. However, because neither the respondents from the *questionnaire survey*, nor the participants from the observed *Gutenberg Project* recognised monetary reward as their primary motivation, this study does not investigate how remuneration influences participation.

The '*accessibility of peer contributions*' is closely related to the process of crowdsourcing:

- 'View' signifies that the contributions are publicly available to any potential contributor; for example, for the application of crowd rating.
- 'Access' signifies that participants are not only able to view others' contributions, but they can also comment on them. This is useful when the crowdsourcer wants to collect the public's opinion on the crowdsourcing outputs; for example, in crowd creation.
- 'Modify' indicates that participants can edit others' contributions, or even delete them. It is often used in crowd processing, where the participants are expected to improve or enhance the outputs by themselves, for example in the case of *Wikipedia*.
- 'None' indicates that an individual's contribution is isolated from the others' and no means of interaction is included.

The setup of the dimension of 'accessibility' in the crowdsourced environment depends on how the crowdsourcer treats the crowd's inputs: whether a collective result or diverse contributions are needed. In my research, collaborative translation projects in the *Yeeyan* community are set up with a high level of accessibility (see section 3.2.1). The quality assessment in 6.4

find that high accessibility of peer contributions plays a significant role in ensuring the translation quality.

The system's perspective

The system's perspective of crowdsourcing investigates the infrastructure of crowdsourcing platforms, and is present mainly in information system studies (Zhao and Zhu, 2012). It is closely connected with the organisation perspective. The crowdsourcer can develop their own system or use third-party systems according to their initiative of crowdsourcing and the design of the process. Technical support of dynamic participation and the interaction between participants, or the participants and the crowdsourcer, is the basic consideration in designing crowdsourcing systems (Vukovic, 2009). This study does not investigate the technical side of crowdsourcing systems, although the collaborative affordances of the *Yeeyan* setup are explained in sections 3.2.2 and 5.1.

Zhao and Zhu (2012) also categorised the design of incentive instruments into the system's perspective, because of the growing attention paid to them recently. For example, certain scholars have been exploring incentive mechanisms from a game theory perspective (O'Hagan and Mangiron, 2013; Packham, 2016). Incentives are also linked to the user's perspective, because crowdsourcers need to understand the full picture of the crowd's motivations to participate in order to incorporate the appropriate motivational elements in the design of incentive systems. My research contributes to the fuller understanding of the motivations and influential factors that may encourage and de-motivate people's participation by analysing in detail the setup, management and output quality of several crowdsourced projects conducted on the *Yeeyan* platform.

2.2 Crowdsourcing Translation

Various terms are proposed to refer to the new translation practice of inviting the general public to translate.

- The *Common Sense Advisory*, a market research company specialising in surveying the Language Services Industry, used the term 'CT3' which stands for 'community, crowdsourced and collaborative' translation (DePalma and Kelly, 2008).
- O'Hagan (2009:97) put forward 'user-generated translation' to indicate translation and localisation 'undertaken by unspecified self-selected individuals' and 'carried out based on free user participation in digital media spaces'.
- Perrino (2009:62) also used 'user-generated translation' to indicate 'the harnessing of Web 2.0 services and tools to make online content – be it written, audio or video – accessible in a variety of languages'. Perrino (ibid) further specified that user-generated translation 'requires human expertise and implies the collaboration between users – be they amateurs or experts'.
- Cronin (2010) used 'open translation' to refer to online translation projects, such as *Project Lingua* in *Global Voices*[3], *Wiki Project Echo*[4], and *TED Open Translation Project*[5], and emphasised 'the subversive potential of the crowd' in framing crowdsourcing practices.
- Pym (2011) recommended the terms – 'community translation', 'volunteer translation' and 'crowdsourcing and collaborative translation', which contain the key features of the novel participatory practices.
- O'Hagan (2011:14) indicated that the lack of monetary reward is

[3] Global Voices is an international community of writers, bloggers and digital activists that aim to translate and report on what is being said in citizen media worldwide. Details see https://globalvoices.org/.

[4] Wiki Project Echo is the collaborative project on Wikipedia which mainly imports and translates content from foreign language Wikipedias.

[5] TED Open Translation Project relies on volunteers to subtitle TED Talks to reach the global audiences. Details see https://www.ted.com/participate/translate.

implied in ‘volunteer translation’, although it also describes the ‘self-initiated action typical of community translation’.

- ‘Non-professional translation’ is also a term used by scholars (Dolmaya, 2012; Pérez-González and Susam-Saraeva, 2012). However, the modifier of the term contradicts the ‘Who’ factor of crowdsourcing, as in fact both professional translators and non-professional amateurs can contribute to crowdsourcing activities.

In order to unify the variety of terms and definitions proposed so far – some of which were mentioned above – O’Hagan (2011:14) extracted the key common characteristics shared in the above terms and highlighted that *crowdsourced translation* is the type of ‘translation performed voluntarily by internet users and is usually produced in some form of collaboration often on specific platforms by a group of people forming an online community’. Jiménez-Crespo (2017) favoured ‘crowdsourcing’ and ‘online collaborative translation’ after reviewing the existing studies and the specific characteristics of these two interrelated practices. He defined the notion of *crowdsourcing translation* as:

> ‘*Collaborative translation processes performed through dedicated web platforms that are initiated by companies or organisations and in which participants collaborate with motivations other than strictly monetary*.’ (Jiménez-Crespo, 2017:25)

Meanwhile, he also defined *online collaborative translation* as follows:

> ‘*Collaborative translation processes in the web initiated by self-organised online communities in which participants collaborate with motivations other than monetary*.’ (ibid)

The difference between these two terms lies in the initiator of the translation process: in the first case it is a company or organisation, and in the other it is a self-organised online community.

In my research, the specificities of the *Yeeyan* community which essentially acts as both translation requester and self-organised community result in no need for a distinction between crowdsourcing translation and online

collaborative translation in terms of the initiator. Furthermore, my investigation focuses on 'collaborative' translation projects in which each individual performs a fraction of the translation-related activity.

Jiménez-Crespo (2017) mapped the concepts related to 'crowdsourcing' in Translation Studies to show the conceptual overlap of the terms related to translational practices that occur both with and without web-mediation (Figure 4). Under the umbrella of *social translation*, *community translation* is seen as the translation performed to facilitate public service that takes place outside the web environment. *User-generated translation* is the most generic term, in which the translators are mostly end-users or readers of the translation.

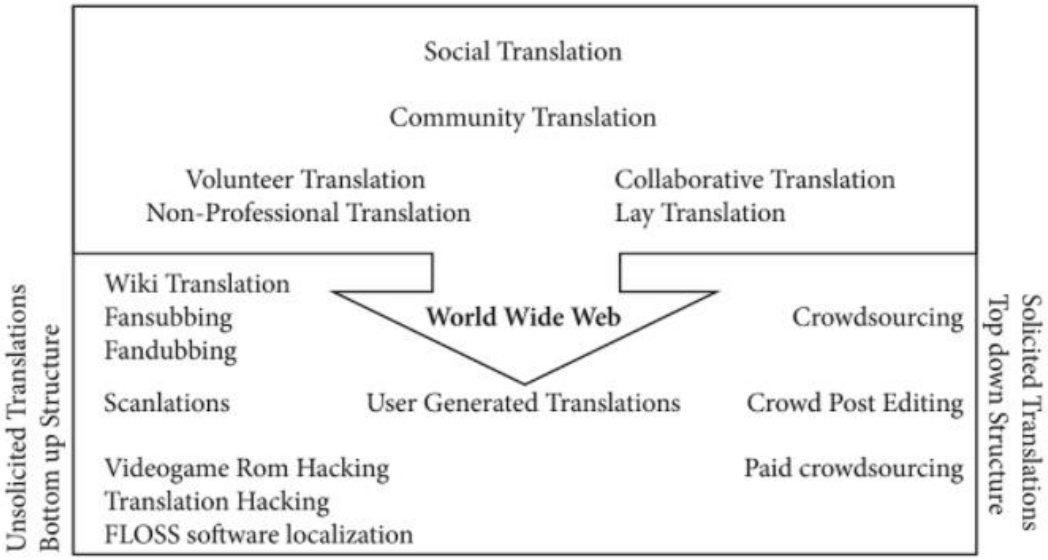

Figure 4. Mapping concepts related to crowdsourcing in Translation Studies (Jiménez-Crespo, 2017:28)

Figure 4 then represents how all web-based translations fall into one of the two models, depending whether they are 'solicited' or 'unsolicited'. According to O'Hagan (2013), solicited translation refers to the cases in which a company, institution or non-profit organisation initiates the call to a community for a translation/localisation task, such as *Facebook*, *Skype*, *Adobe*, and *Mozilla Firefox* (Jiménez-Crespo, 2011, 2013b; Mesipuu, 2012)

or NGOs such as *The Rosetta Foundation*[6] (Anastasiou and Schäler, 2010), and *Kiva*[7] (Baer and Nagle, 2011). It matches Jiménez-Crespo's (2017) definition of crowdsourcing translation with regard to the initiator.

On the other hand, *unsolicited translation* refers to the cases in which self-selected collectives of users undertake self-organised specific translation tasks without responding to any specific request. It fits Jiménez-Crespo's (2017) definition of online collaborative translation. Audio-visual collaborative practices, such as fansubbing, fan dubbing, scanlations, and videogame romhacking are typical examples of such kind. Their translations are usually distributed and shared via websites, blogs or any other type of digital media. The localisation of free open source software (FOSS) falls into this category, as well.

Jiménez-Crespo's and O'Hagan's classifications, however, differ based on a traditional perceived value of the translation requester, assigning the needs of a community of users to the 'unsolicited' category. This classification simply reinforces traditional methods and is not always mirrored by the impact and distribution of the translated product – 'unsolicited' translations can have much larger distributions than 'solicited' ones, yet traditional practices can view them as inferior in quality due to the absence of one single, established and recognised requester.

My research will demonstrate that, even though at first sight a large part of the *Yeeyan* activity can be considered 'unsolicited', 'wiki translation'-similar, its practices, organisation and output quality are not dissimilar to those of 'solicited', 'crowdsourcing' translation. I therefore propose that the distinctions made in Figure 4 and the allocation of 'crowdsourcing' at one end of the spectrum are rather artificial, and the reality is much more homogeneous.

---

[6] The Rosetta Foundation is a non-profit organisation registered in Ireland promoting equal access to information and knowledge across the languages of the world. Details see https://www.therosettafoundation.org

[7] Kiva is an international non-profit organisation based in San Francisco, with a mission to connect people through lending to alleviate poverty. Kiva's Review and Translation Program relies on volunteers to translate local profiles and review them to ensure that each profile is understandable to lenders and complies with Kiva's policies. Details see https://www.kiva.org/work-with-us/reviewers/translationprogram

One may even argue that such detailed distinctions are purely academic exercises, because the industry tends to favour a much simpler and coherent classifications.

2.2.1 Classifications of crowdsourcing translation

It has been observed that there is a diversity of activities that fall under the umbrella ‘crowdsourcing translation’. The general phenomenon of crowdsourcing translation practices can be described as: ‘small tasks carried out individually by collaborators whose work may or may not be pooled’ (Morera-Mesa, 2014:33), for example in *Amazon Mechanical Turk* and *Facebook*. Small tasks refer to the small translation units – often single sentences or phrases – which are done by individual users; and ‘bigger tasks that are solved by a number of members of the crowd that collaborate’ (ibid), for example, the drafting and translation of a *Wikipedia* entry, blog translation in *Global Voices*. Bigger tasks refer to large translation units (e.g. a paragraph of the text or the whole text) which are done either by individual users or several users in a collaborative manner. As an information-oriented community, *Yeeyan* belongs to the latter category, using the crowd to perform ‘bigger tasks’.

In addition to solicited and unsolicited models, there are also further classifications of crowdsourcing translation. With regard to the type of initiative, Bey et al., (2006) divided it into *mission-oriented* and *subject-oriented*. The former refers to strongly-coordinated groups of volunteers that are involved in translating clearly-defined sets of documents – e.g. technical documentation of projects like *Mozilla* (Mozilla, 2017). If one also adopts the different implied status of the translation requester behind the distinction between ‘solicited’ and ‘unsolicited’ division, mission-oriented can be mapped onto O’Hagan’s classification of *solicited crowdsourcing*. The latter refers to individual translators who translate online documents such as news and reports and make translations available on personal or group web pages. These groups of translators are not assumed to have any orientation and training in advance, but they share similar opinions about events such as humanitarian help, news and report translation. It can be mapped into

O'Hagan's classification of unsolicited crowdsourcing.

DePalma and Kelly (2011) categorised crowdsourcing practices according to three different environments: *cause-driven*, *product-driven*, and *outsourced-driven*.

- The first type of crowdsourcing is driven by a cause, meaning that people choose content that interests them. Collaborative translation is produced by people who share a specific common goal, such as offering help in the wake of disasters, as happened after the Haiti earthquake (Munro, 2010) and *Yeeyan's* collaborative translation project on the earthquake safety manual after the Wenchuan earthquake (Yeeyan, 2008). Volunteer translation in the *Rosetta Foundation*, *Translation Commons*, or *Translators without Borders* are other typical examples of such a type of crowdsourcing. Volunteers are generally not remunerated. Media content is a popular cause-driven source of crowdsourced translation, as well (Kelly et al., 2011). *Fansubbing* and *fan translation* also fall into the cause-driven category.
- The second type, product-driven crowdsourcing refers to projects initiated and managed by for-profit companies, which belongs to the solicited model of crowdsourcing. Companies such as *Adobe* (Ray and Kelly, 2011) have adopted product-driven crowdsourcing for software localisation. Volunteers usually gain some kind of remuneration from the company, e.g. giveaways of free products/services, promotional merchandise from the company, or other intangible forms (reward on the community boards, praise or recognition within the community) (Kelly et al., 2011).
- The third category refers to outsourcing portals that offer crowdsourcing translation 'as either their revenue driver or as one of their service offerings[8]' (ibid:90). People who perform language-

[8] Examples of outsourcing portals that offer crowdsourcing translation service: Crowdflower (http://crowdflower.com/) , Microtask(http://www.microtask.com), OneHourTranslation (http://www.onehourtranslation.com), Amazon Mechanical Turk (https://www.mturk.com), Zhubajie (http://www.zhubajie.com), and Zuodao (https://www.zuodao.com).

related activities through such portals gain monetary remuneration.

Moreover, Mesipuu (2012) divided crowdsourcing translation according to the forms of participation as: *open* model – any member of a community can openly participate (e.g. *Facebook*); and *closed* model – only the preselected or invited members by the crowdsourcer can participate (e.g. *Kiva*, *Skype*). The classification relates to Geiger et al.'s (2011) meta dimensions of crowdsourcing organisation: 'preselection of contributors' and 'accessibility of peer contributions'.

The subject of my research – *Yeeyan* belongs to the category of using the crowd to perform 'bigger tasks' in terms of the translation unit. It is a *cause-driven* community which endeavours to disseminate foreign cultural content. At the same time, it is also *product-driven* because of its book translation and publication project. As for the process of crowdsourcing, both *open* and *closed* models are implemented depending on the purpose and quality requirements of the project. Hence, this study is expected to provide a more thorough evaluation of the crowdsourcing model as the *Yeeyan* community encompasses a combination of the features exhibited in the existing crowdsourcing initiatives while also having its own characteristics.

2.2.2 What makes crowdsourcing translation possible?

2.2.2.1 Translation as a natural human ability

One of the substantial differences between crowdsourcing translation and professional translation is represented by the actual operator of the translation activity. This is also the major criticism of crowdsourcing (Désilets and van der Meer, 2011; Ellis, 2009): how can a group of people – some with, some without formally-validated translation competences – perform such a linguistically-sophisticated activity?

Translation Studies is a discipline that focuses traditionally on professional translation as a complex, specialised skill (Whyatt, 2017). It is only in recent years that the translation and other related services provided by non-professional translators have begun to attract more attention (Pérez-González and Susam-Saraeva, 2012).

Regarding the problem of ‘who can translate’, scholars hold different views as one side claims that special qualification is needed, while the other argues that it is possible for people with bilingual skills, be they natural bilinguals (people raised with two languages and two cultures) or L2 learners, to perform well as translators.

Malakoff and Hakuta (1991:144) supported the former view, that translation is a valuable skill that is available only to the ‘highly trained and linguistically sophisticated bilinguals who come out of interpreter and translator training schools’. Interlingual translation or ‘translation proper’ is classified as a special skill by De Groot (1997) who further claimed that it is a privilege of the few rather than a general aptitude.

However, Harris (1977) put forward the idea of ‘natural translation’ and explained that the human ability to translate is not a special skill that requires formal training, but it is available even to bilingual children (Harris and Sherwood, 1978). By highlighting the importance of natural translation, Harris (1977) stressed that it deserves to be researched within the new discipline of Translation Studies. Yet, Harris’s view was heavily criticised at the time and was considered as marginal in Translation Studies (Whyatt, 2017).

The theory of natural translation has been gaining ground in recent years. Whyatt (2017) underlined the communicative nature of human beings in conjunction with the communicative function of translation. With the acceleration of globalisation, intensive migration has formed multilingual and multicultural communities, which has naturally led to the increasing need for interlingual and intercultural mediators. However, because these intercultural communicative occasions usually occur in informal settings, professional translation services are not always involved in such interactions, such as shopping events, hospital visits, police intervention, immigration policies, etc.; in addition, because of economic austerity, those taking part in these situations are at times left without the assistance of professional mediators (Meyer et al. 2010). In these cases, bilinguals in multicultural communities play an effective role in facilitating communication. As Whyatt (2017:48) put it, ‘non-professional translation makes room for a reassessment of the human

ability to translate as a part of intellectual potential'.

Whyatt (2012:29) also proposed a developmental continuum of the human ability to translate, in which translation 'ability' is seen as the initial stage that is open to all bilinguals. The term 'bilingual' is used here in its liberal sense following the psycholinguistic tradition (Bialystok, 2009), including anyone who can communicate in two different languages. The *developmental perspective* views non-professional translation as the predisposition, described by Harris and Sherwood (1978), to professional expertise. The model agrees with Shreve's view (1997:125) that translation ability should be seen in 'a kind of evolutionary space', in which the natural ability of bilinguals to translate is only the starting point. Translation ability can evolve to the ultimate stage under favourable external circumstances and internal conditions. To be more specific, the need for translation services and the availability of adequate training opportunities are among the external conditions, while the translator's intentional effort to develop is the internal condition. The combination drives the evolution of the human ability to translate (Whyatt, 2012).

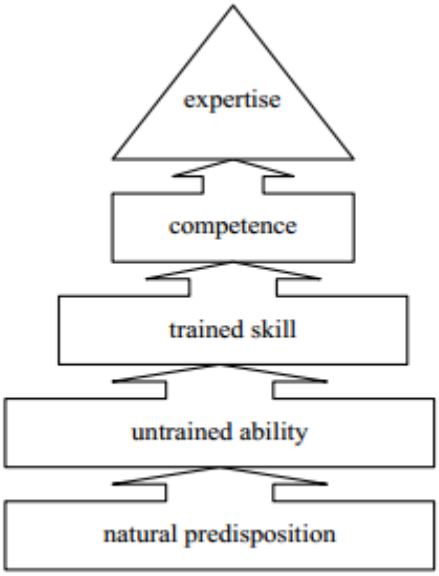

Figure 5. Evolution of translation as a human skill (Whyatt, 2012:29)

As shown in Figure 5, and to some extent mirroring the human natural ability to communicate, one could argue that all people who have access to at least

two languages are predisposed to translate if the need arises. As Whyatt (2017:61) puts:

> *'The capacity to act as an intercultural mediator in a communicative event is developmental in nature and ranges from the untrained ability of natural translators, which can develop into a more refined skill if the experience of translation is treated as an educational learning environment by the translating individual. If the trained skill is utilised and coupled with self-development aimed at improved translation performance, professional competence is gradually acquired, which in some cases can result in expert performance.'*

It is worth pointing out that in the developmental continuum, only some of the natural translators will make the effort to refine their ability and become capable of expert performance. It depends on whether the person chooses to pursue a career in translation, either with or without the support of structured education (Hoffman, 1997; Whyatt, 2017).

According to the latest UK Translator Survey (2017) conducted by *the European Commission Representation* in the UK, *the Chartered Institute of Linguists* (CIOL) and *the Institute of Translation and Interpreting* (ITI), out of 588 respondents, 24% do not have a formal qualification in translation but are actively involved in the industry as professional translators. The figure reinforces the idea of natural translation in a sense that a special qualification is not the prerequisite. It reflects L2 learners' potential for performing translation activities or pursuing a career in translation.

In a crowdsourced environment, it is not always clear whether each translator is professional or non-professional. Nevertheless, they must all be L2 learners of varying proficiencies in order to access the predisposed ability to translate. The idea of 'natural translation' and Whyatt's (2012, 2017) developmental continuum provide a reasonable explanation for why crowdsourcing translation is possible.

2.2.2.1.1 Limited capacity of natural translation

Several studies have already investigated people's ability for natural translation in the last two decades and have found that bilingual children can both interpret and translate materials that are within their comprehension and vocabulary (Harris, 1980; Malakoff, 1992). As for the quality of the translation, Malakoff and Hakuta (1991:161) found that natural translators tend to focus on the meaning of the message that needs to be conveyed rather than on the linguistic norm, that is 'when the quality of the translation suffers, the errors are usually in sentence structure and not in meaning'. The *conceptual* level (semantic, relating to the meaning of the message) is maintained and delivered, while the *structural* (lexico-grammatical) level may be compromised during natural translation (Whyatt, 2017). L2 learners, similar to bilingual children, can all be expected to act as natural translators, at least initially, given that they represent 'the type of translators perceived as amateur and unskilled, who by linear transcoding produce awkward, incomprehensible translations full of L1 interference (i.e. translationese)' (ibid:52). The same performance is also observed by Gile (2004) in trainees who can nevertheless be taught to translate in a sense-oriented approach, thus addressing the problem of 'translationese'.

The wide practices of crowdsourcing translation (e.g. *Facebook*, *Global Voices*, *TE*D, *Mozilla Firefox*, etc.) have already demonstrated the human ability for natural translation. Yet, there is little empirical research that measures the individual differences in terms of translation competence, identifying the stage people are at in the developmental continuum, thus aiding the effective planning of the crowdsourcing process. Moreover, there is also a shortage of longitudinal studies that explore the evolutionary perspective of translation ability through continuous practice.

Therefore, in this PhD project, inspired by the idea of 'natural translation', I examine the translation ability of crowdsourcing participants from two perspectives: first, do translators in the crowdsourced environment have a similar level of translation ability as natural bilinguals in previous studies, and do they focus more on rendering into the target language the semantic

features of the source text while at the same time neglecting the syntactic structure? Secondly, is there a developmental process during people's participation in crowdsourcing translation?

2.2.2.2 Participatory culture: harnessing collective intelligence through the discourse of Web 2.0

'Natural translation' explains the primary factor that makes crowdsourcing translation possible from the perspective of human competency. This section reviews the technological and cultural factors that support the functioning of crowdsourcing.

The development of information and communication technology has changed the methods of content generation. As part of Web 2.0 advancements, a number of web applications have been created, among which user-generated content (UGC) platforms are the most prominent ones that pushed the role of 'user' under the spotlight (Dijck, 2009). *YouTube* and *Wikipedia* are two well-known examples of UGC platforms. People of all ages and backgrounds are increasingly engaged in network activities. They can create and distribute content via online and interpersonal networks at an ever-accelerating rate. Citizens are provided with inexpensive tools to capture, edit and organise real-time data and multimedia content to promote personal and political interests, which create a new cultural sphere of participation via the internet. Web 2.0 represents the support that makes such participation feasible (Waldron, 2013).

Web 2.0 emphasises a high level of usability, interoperability, interactivity and collaboration, leading to the significant growth of online communities that gather people with similar interests and preferences to create and distribute content as the community sees fit (Dijck, 2009; Waldron, 2013). The once professional-dominated industries, particularly in the media domain, have witnessed a profound and prolonged transition (Jenkins and Deuze, 2008). People that used to be the passive recipients of earlier stages of media content (*content-consumer*) have switched their role to active producers or creators (*content-prosumer*) (Dijck, 2009; Jenkins and Deuze, 2008; Waldron, 2013; Zhang and Mao, 2013). They are the main force that

contributes to the vast amount of UGC on the internet (Dijck, 2009).

Cultural theorist Henry Jenkins (2006) saw the paradigm shift in the way media content is produced and circulated, and put forward the idea of 'Participatory Culture'. It is defined as the one with

> *'1. relatively low barriers to artistic expression and civic engagement, 2. strong support for creating and sharing creations with others, 3. some type of informal mentorship whereby what is known by the most experienced is passed along to novices, 4. members who believe that their contributions matter, and 5. members who feel some degree of social connection with one another (at the least, they care what other people think about what they have created)'* (Jenkins et al., 2009:5-6).

Though Participatory Culture is usually used to discuss the new phenomenon of media content creation and consumption given the advent of web technology, it also offers a theoretical explanation of the proliferation of online translation communities, such as fan translation and crowdsourcing translation. The key elements in its definition are in accord with the principles followed by online translation communities: low barriers to engagement, support for content creation, and sharing.

Taking crowdsourcing translation as an example, its openness encourages a wide range of participation within the community, thus, individuals can get engaged as translators, revisers, active readers, commentators, or even multiple-role players. All kinds of contributions are valued in supporting the functioning and future development of the community: as far as translators are concerned, they are the major force in the production of the translation work; revisers are those who improve the quality of translations; active readers are consumers of the translation products; and commentators are those who provide feedback that encourages other kinds of participation. The collective work of all these functioning roles accumulates to achieve a common goal.

Participatory Culture manifests the phenomenon that the philosopher Pierre Lévy (1999) called the emergence of collective intelligence – the effect of

accumulating diverse participation in which every individual's expertise is valued (Delwiche and Henderson, 2013). This phenomenon is very relevant in crowdsourcing translation, as it stresses both the diversity of participation and the value of each individual's contribution.

2.2.2.2.1 Participatory Culture from a 'user' perspective – community building

The current trend of participatory culture in content generation is theorised by emphasising the users' strong preference to share knowledge and culture in communities (Dijck, 2009). Scholars tend to highlight the 'user' perspective when explaining how a cohort of people gather and collaborate to create something meaningful to the group.

To begin with, the 'user' generally represents the active internet contributors, who put in a 'certain amount of creative effort' which is 'created outside of professional routines and platforms' (Livingstone, 2004). Moreover, 'community', in relation to UGC sites and user groups signifies the 'inclination of users to belong to a (real-life) group and be involved in a common cause (Dijck, 2009). The concept of 'community' connects groups with common preferences such as music, movies or books (a so-called 'taste community'); building 'taste' is an activity that necessarily connects individuals with social groups (Hennion, 2007). For instance, fan translation for anime and games, or fan-subbing for different categories of videos, are the result of the collective effort made by 'users' with the same taste. The apparent reason for such participation is because of 'personal interest', yet there are more motivations involved (for a detailed discussion, see section 2.3).

Different from the product-oriented or process-oriented research paradigms in traditional Translation Studies, O'Hagan and Mangiron (2013:277) put forward a new research direction: to 'explore a research impetus which is coming from user empowerment afforded by new communications environments, promoting collaboration and sharing among users' to emphasise the role of fan activity in game localisation, highlighting the importance of 'users' and teamwork in fan translation work. In the discussion of fan work, O'Hagan and Mangiron (2013) indicated that striving to preserve

the authenticity of the original product in translation and localisation is what drives the mass phenomenon of fan translation in the game industry. It can be considered as another important factor that attracts participation from the user perspective. Similar to fan translation, translators in crowdsourced environments are more likely to understand the source context and the target culture, and to know what the target readers want or respond to.

Participants – or users – are the fundamental force behind crowdsourcing projects, and they are grouped into several categories according to the importance of their contributions, as discussed in section 2.1.2.2. The significance of *lead* users in Rajala et al.'s (2013) study is also echoed in fan translation work (O'Hagan and Mangiron, 2013). Since lead users usually take part in highly-skilled activities, their work is well-respected in the community and fosters a recognisable status within the community. Taking *Yeeyan* as an example, one's status can be inferred from the number of followers listed in their user profile: those who produce a large number of translations and actively give others feedback are likely to have more followers (fans)**.** This aspect matches the developmental continuum proposed by Whyatt (2012, 2017), as well.

2.2.2.2.2 Participatory Culture from a 'business' perspective

Dijck (2009) indicates that the full implementation of Web 2.0 and the boom of UGC sites has also influenced the business interest, which has shifted away from consuming activities and gravitated towards producing activities, giving users more power over content because they add business value. Participatory Culture enhances the potential for *niche* marketing, as advanced digital technologies facilitate the tracking of individual social behaviour through user engagement. A closer relationship between content producer, advertiser and consumer is established since 'specific people consuming specific content constitutes the basis of targeted advertising; information vital to marketers derives from the connection between content, disposable income and consuming behaviour' (ibid:47). *Facebook* is a good example of this kind. Crowdsourcing translation attracts those who are interested in source content (Yunker, 2014), which facilitates the expansion

of a niche market of specific content and service.

Apart from tracking the users' preference for content, products and services, another benefit is the user's new role under Participatory Culture: that of *content* and *data provider*. By creating and uploading content on UGC platforms, users also (sometimes willingly and unknowingly) provide information about their profile and activity to site owners and metadata aggregators, such as gender, age, nationality, income, etc. The service agreement of a site (i.e. Terms of Use) usually grants permission for the site owner to use such metadata. Metadata can be mined for various purposes, from targeted advertising to interface optimisation, which are all beneficial from a business perspective (Dijck, 2009). All UGC users, whether active creators or passive viewers, form an attractive demographic to advertisers on UGC platforms (Li et al., 2007).

Crowdsourcing translation, especially applications that aim at the localisation of products for commercial purposes, meets the aforementioned benefits from the business perspective. Besides, the growing size of UGC shortens the shelf-life of texts on the web, which makes it economically challenging to keep recruiting and paying for translating content of uncertain popularity and circulation. Crowdsourcing, from this standpoint, appears to be a relatively feasible approach to address the huge amount of translation work.

2.2.2.3 Summary

To summarise, this section reviewed the factors that make crowdsourcing translation possible. It began with the agents of translation – arguing that translation is a human natural ability for bilinguals, with the skill level of the user progressing along a developmental continuum. The second part provided explanations from the technological and cultural perspectives. Web 2.0 enables the organisation and participation of crowdsourcing activities, and also cultivates a Participatory Culture which theoretically explains user engagement. The 'user' perspective underlines the features that attract people's participation; the 'business' perspective highlights the benefits of enhanced user empowerment.

2.2.3 Benefits of crowdsourcing in the context of translation

In addition to the ‘user’ and ‘business’ perspectives of participating and adopting crowdsourcing practices in section 2.2.2.2, many benefits have been attributed to using crowdsourcing in the context of translation.

To begin with, the most obvious advantage in deploying crowdsourcing translation is quicker solution with lower (sometimes free) cost (Sajan, 2013). As crowdsourcing translation is usually done for no or little payment, people who sign up for the projects are mainly volunteers, which saves the labour cost (Ma, 2017). Despite the preparation and implementation of crowdsourcing, the actual time for a translation task to be carried out is very short, especially for ‘small tasks’ which may take only minutes to be carried out. This is because the ‘open’ call in crowdsourcing and many people have access to the task (Anastasiou and Gupta, 2011).

Secondly, crowdsourcing alleviates the imbalance between major and minor languages in the Language Services Industry (Désilets and van der Meer, 2011). Taking localisation as an example, in some cases it is not feasible to make minor languages available through traditional processes because they may not have a strong enough market impact (Filip, 2012). However, minor-language users can easily engage in language-related activities through crowdsourcing, which on the one hand reduces the cost of language services, and on the other hand increases the global reach of their beneficiaries. Post et al. (2012) constructed parallel corpora for six languages from the Indian subcontinents using crowdsourcing, which is expected to support the development of machine translation of under-resourced and understudied languages.

Thirdly, crowdsourcing addresses a genuine need by offering translations for UGC which is not included in the normal market share of commercial translation companies. In addition, crowdsourcing translation processes can also result in a faster turnaround (Lommel and Ray, 2009). Thus, crowdsourcing appears highly suitable for the translation of UGC which is considered as transient content (Désilets and van der Meer, 2011).

Another benefit comes from the specialist knowledge of users (Tsang, 2008). As O'Hagan (2011) indicated, some of the contributors are domain-professionals or activists on a project that they are highly interested in or they consider worthwhile even if there is no remuneration. The involvement of such expert users who sometimes work on several projects or different parts of the same project can lead to quick translation or localisation solutions which in turn contribute to a faster turnaround (Lommel and Ray, 2009; Kelly et al., 2011).

The potential to increase brand loyalty is another claimed benefit of crowdsourcing translation (Désilets and van der Meer, 2011). Because it is a community-based activity which enhances the user engagement with content/product, people who have participated in crowdsourcing translation are expected to devote more to the community and the crowdsourcer.

Moreover, scholars also highlighted quality-related benefits such as the production of more native-sounding translations through crowdsourcing (Désilets and van der Meer, 2011; Meyer, 2011). As discussed in section 2.2.2.2.1, crowdsourcing strongly relies on community building. The shared linguistic resources within a community can help improve the quality of the translation and the benefits of sharing resources can be spread more easily and faster in a community environment (Désilets and van der Meer, 2011). Higher quality of translation here refers to 'the acceptability of the translation among its final users' (Morera-Mesa, 2014:47). Jiménez-Crespo (2013a) empirically tested the naturalness of crowdsourcing translation by comparing the output of Spanish translations in *Facebook* with native Spanish social networking sites. The result showed that the localised version of *Facebook* was similarly fluent and natural when compared to natively-designed Spanish social networking sites, including the use of the most frequent Spanish lexical units and syntactic patterns.

Another quality-related benefit results from the active role of expert participants in the crowdsourced environment. It is argued that if professional translators/reviewers or expert users are involved in the revision process as gatekeepers and overseers, the quality of the translation can be improved by

these experts who will correct macro-textual errors like inconsistent terminology (Jiménez-Crespo, 2011). Kelly et al. (2011) further highlighted the feedback benefit provided by these experts. As feedback can be delivered faster in the community environment, it can result in the prevention of potential errors, thus generating higher quality.

Finally, crowdsourcing is seen as a good learning environment for trainee translators and novices (Dolmaya, 2011). Jiménez-Crespo (2017) discussed how peer feedback in the online collaborative environment facilitates learning from a socio-constructivist perspective and proposed to use a crowdsourcing environment for translation training (for more details, please see section 2.2.5).

2.2.4 Criticisms of crowdsourcing in the context of translation

Despite the already-mentioned advantages of crowdsourcing, the research literature also includes several criticisms. Dolmaya (2011), for instance, criticised the increased 'visibility' of translators achieved through the reward mechanisms in the crowdsourced environment. In some situations, crowdsourcing translations are actually revised by professionals, but often the amateurs gain visibility, which can negatively influence the public's perception of the value of translation.

Moreover, the growing popularity of crowdsourcing translation has been seen as a threat by professionals. In 2009, *Translators for Ethical Business Practices* created its first online petition against crowdsourcing on *Facebook* and *Twitter* (Proz, 2009). However, scholars believe that crowdsourcing translation will not destroy the professional market, in the same way that Free Open Source Software has not destroyed the proprietary software market (Désilets, 2007; Kelly et al., 2011).

Furthermore, it is also suggested that the initial advertising/marketing effect of crowdsourcing may not work as well as expected (Ray and Kelly, 2011). If approached in the wrong way, a 'backlash' effect may be generated within the crowdsourcing translators community – for example, the failure to launch crowdsourcing translation in *LinkedIn* (Kelly, 2009).

Other criticisms are related to the quality of crowdsourcing translation (Désilets and van der Meer, 2011; Ellis, 2009). To begin with, some people participate in crowdsourcing activities wishing to help in the case of disasters. However, such emotional attachment does not guarantee translation quality (Morera-Mesa, 2014). Also, malicious users and users that are motivated by material rewards, as well as well-meaning contributors with poor linguistic skills can also produce poor translations (Zaidan and Callison-Burch, 2011). It is difficult to control quality with a vast pool of unknown translators.

Parallelism can be a disadvantage in terms of consistency. Yahaya (2008) observed that the more translators are engaged in a project, the higher the chance of inconsistencies in the translation. Munro et al. (2010) proposed to reduce the negative impact of having large number of translators by promoting the communication among them. They suggested to implement fora and chat rooms to help deal with inconsistency issues. Empirical data in my research shows the extent to which collaborative resources/tools – i.e. shared glossary lists and peer-to-peer interactions in the chat board – help increase consistency (discussed in Chapter 6 and 6.4).

The technical settings of a crowdsourcing project can also impact on the quality of the product. It is argued that small tasks in which the translation unit is a *string* de-contextualise the source text (Morera-Mesa, 2014). Such settings may hinder translation quality, as individuals are exposed to small pieces of the text at a time – e.g. one sentence or one phrase – which influences their understanding of the source content. Georgakopoulou and Sosoni (2016) conducted an experiment with trainees to imitate using crowdsourcing for parallel translation corpora creation. The result showed that translating large units (e.g. paragraph or whole text) achieved higher accuracy than working on smaller units. By contrast, Jiménez-Crespo (2011) supported the approach used by *Facebook*, indicating that short, isolated strings are more suitable for the workflow of using multiple translations and letting the users vote for the best one.

Finally, Désilets (2007) indicated that deadlines do not work well for crowdsourcing translation: the lack of a contract makes it difficult to enforce a

deadline. He observed that expecting the entire translation to be completed before the deadline is not always going to be possible.

2.2.5 The learning perspective of crowdsourcing translation

As discussed in the developmental continuum of human translation competence and the benefits of crowdsourcing translation, academics are intrigued by the potential of learning in the crowdsourced environment. O'Hagan (2008:178) put forward the idea of fan translation being an 'accidental training environment' which seemed to offer 'authentic and situated learning environments for amateur translators who are well motivated'. Jiménez-Crespo (2017:238) recommended the introduction of crowdsourcing and online collaboration in institutionalised learning as 'grounds for playful experimentation for future professionals'. Online communities provide diverse translation possibilities through which amateurs or trainee translators can engage in hands-on learning with the expanded source contexts, and learn in 'casual, relaxed and self-directed ways' (ibid:227).

'Deliberate practices' and 'feedback' are considered the components that promote the development of higher competence in the crowdsourced environment. It is argued that 'if the acquisition of translation competence in institutions of higher learning is considered as the acquisition of "expert knowledge" (Shreve, 2006) that develops through "deliberate practice", then any guided and structured intensive "deliberate practice" setting, with different levels of difficulty and appropriate feedback to develop expertise (Ericsson, 2000) should be mostly welcomed by the training and professional community' (Jiménez-Crespo, 2017:231).

Given the diverse industry-informed applications, crowdsourcing projects appear to be a suitable learning environment to enhance translation competence acquisition, due to the different roles and tasks fulfilled during the users' participation. In a fast-evolving localisation industry, adopting an instructor-led rather than a facilitator-led training approach (Dobson in Schwimmer, 2017:52) is detrimental to trainee linguists because it does not

allow them to experience genuine collaboration, alongside a comprehensive range of multi-faceted challenges present in a complete translation/localisation workflow.

Socio-constructivist approaches argued that collaborative translation projects based on real-world scenarios are the ideal environments for 'situated learning in a semi-professional setting' (Jiménez-Crespo, 2017:232). The progression from instructor-centred learning to interactive, formative, collaborative 'socio-personal' approaches benefits from using crowdsourcing for learning, due to the environment's unique blend of autonomy and real-world collaborative practice (Kiraly, 2000).

Scholars have also pointed out that online learning contexts do not necessarily need the instructor to provide feedback, as peers can be a good source of feedback (Kiraly, 2000; Massey and Brändli, 2016). Neunzig and Tanqueiro (2005) classified the possible formative feedback in online settings as:

- 'Anticipatory feedback': feedback that is provided when the 'translation has not yet been formulated in writing' – (ibid: NP) (e.g. providing a glossary or discussing the difficulties in the text).
- Individual feedback: (1) elaborate feedback – feedback provided directly after commission of errors [comprising] prompts or indications of the most appropriate strategy to be applied to find an acceptable solution; (2) simple feedback – an evaluation of the contribution (reinforcement-disapproval), informing him/her whether the answer is correct or incorrect, or it may entail an indication of what the right answer is.

Based on Neunzig and Tanqueiro's (2005) research, Jiménez-Crespo (2017) extended feedback classifications specifically for the crowdsourced environment. 'Individual feedback' can be provided to each participant in different ways. It can appear as 'simple feedback' in which the participant eventually receives a result. It is commonly seen in iterative models or practices with success-based remunerations. Jiménez-Crespo (2017:239)

referred such form as ‘delayed crowd feedback’. The voting mechanism in *Facebook* is a typical example. Though it only provides a minimum level of feedback, it is still considered to be valuable for earlier stages of the learning process (ibid:240).

Elaborate feedback occurs in practices that replicate the traditional TEP (translation – editing – proofreading) approach for projects with integrative contribution and high accessibility of peer contributions. Forums (e.g. *Facebook*) and other independent forums of discussions (e.g. *Proz.com*) between the participants provide elaborate feedback, as well. Under this model, two different types of feedback are proposed: feedback by peers, or feedback by qualified participants. The notion of ‘qualified participant’ refers to ‘those in initiatives that have ranked participants according to any type of criteria’ (ibid:241). It can be based on the evaluation of performance, experience, frequent users, etc.

Non-individual feedback is added to capture the use of translation resources, guidelines, preparatory materials, glossaries, norms and codes in crowdsourcing initiatives. It is a form of *preventive* or *orientative* feedback (ibid).

For initiatives that use crowdsourcing for MT post-editing, and commercial platforms (e.g. *Amazon Mechanical Turk*), no feedback is provided to the participants.

Neunzig and Tanqueiro’s (2005) empirical study concluded that the translation process and the outcome quality partially correlate to (1) the type of feedback employed (whether simple or elaborate), (2) how it is administered (whether individual or ‘ready-made’) and (3) when it is presented (whether anticipatory, contiguous or follow-up). Their result also showed that online elaborate feedback produced the highest acceptance rate among trainee translators. It was the most influential with regard to the modification of translator behaviour and resulted in the highest rate of acceptable translations, which indicated a positive effect on translation quality. In the study of Orrego-Carmona (2014:138), after participation in

volunteer subtitling communities with mentoring programs, 83% of participants thought the communities could be a good source of feedback for their translations. However, the problem is that, so far, few studies have investigated the actual feedback process in crowdsourcing translation. In my research, self-assessment from the *questionnaire survey* and the participants in *Yeeyan*'s *Gutenberg Project* also highlighted that the act of practising translation, combined with receiving peer feedback helped community members improve their translation ability (see section 5.4 in Chapter 5). More importantly, through the observation of actual practices of collaborative translation projects in *Yeeyan*, I collected a substantial amount of interaction data, and I also used *Dialogue Act Analysis* to analyse the giving and receiving of feedback through the question, clarification, discussion, and agreement in conversational discourse (for details see section 6.2.2 in Chapter 6).

In light of the social-constructivist approaches to how crowdsourcing translation can be exploited to improve translation competence, this research examines the learning process in the collaborative projects in *Yeeyan*. It attempts to find out whether learning is gained through participation, the feedback patterns, as well as the roles participants played in the learning process (i.e. whether highly experienced peers or other participants take up the role of instructors as in regular translation training environments).

2.3 Motivations

Several surveys have been conducted on various crowdsourcing websites to investigate the motivations that drive people to participate, covering problem-solving initiatives, open source projects, human intelligence tasking places (e.g. *Amazon Mechanical Turk*), and idea innovation sites (Hars and Ou, 2002; K. R. Lakhani and Wolf, 2005; Zheng et al., 2011). These studies indicated that there are many common reasons why people participate, both intrinsic and extrinsic, but there is no single motivator that applies to all crowdsourcing applications. For instance, the opportunity to develop one's creative skills, build a portfolio for future employment, and challenge oneself to solve a difficult problem are motivators which emerge from several

crowdsourcing cases, but some crowds are driven by financial gain and do not mention these motivators. Drawing from these existing studies, some motivations for individuals in crowds that emerge across more than one case include: ‘To earn money, to develop creative skills, to network with other creative professionals, to build a portfolio for future employment, to challenge oneself to solve a tough problem, to socialise and make friends, to pass the time when bored, to contribute to a large project of common interest, to share with others, and to have fun’ (Brabham, 2013:68). These motivations generally suit the context of crowdsourcing translation. However, they do not encompass all the possible reasons that drive people to invest time and effort to perform the linguistic activity, such as to fulfil political needs, and to effect social change (see the summary of motivation studies in crowdsourcing translation in Table 1).

2.3.1 Sources of motivation

Aforementioned motivating factors that Brabham (ibid) summarised from previous research fall into the motivation theories highlighted in cognitive psychology literature – Cognitive Evaluation Theory (CET), Self Determination Theory (SDT) and General Interest Theory (GIT). These frameworks provide a theoretical foundation which explains how motivations relate to human needs and how they influence the participation in volunteer activities.

The basis of CET is that individuals assess whether the execution of a task meets and fulfils basic psychological needs, focusing on contextual factors of a task, such as monetary rewards, punishments, verbal reinforcements (positive and negative feedback), and deadlines. When contextual factors satisfy a person’s needs, these context factors have a positive effect on the person’s intrinsic motivation and the person will be engaged in this activity. When contextual factors do not satisfy a person’s needs, a negative effect on the individual’s intrinsic motivation will occur, and subsequently it will diminish his or her performance (Deci, 1972). Deci and Ryan (2002) suggested that there are three psychological needs which are relevant in CET: the need for autonomy, the need for competence and the need for social relatedness.

To be more specific, the need for autonomy refers to the extent to which a person him/herself can determine his/her behaviour. It is also recognised in the behavioural economic study that autonomous functioning makes people more engaged and productive (Frey and Stutzer, 2002). According to CET, contextual factors have an effect on the level of autonomy that people experience. Examples of factors that constrain autonomy are social controls, evaluative pressures, rewards, and punishments. Contextual factors that have a positive effect on autonomy are choice and opportunities for self-direction. The need for competence can be understood as the psychological need of an individual for confirmation of one's ability, thus leading to an increase in self-esteem. Factors such as performance feedback and explicit competition which leads to a distinction between winners and losers can enhance or diminish one's perceived competence. The need for social relatedness refers to the need for a warm, caring and secure environment during the execution of the task. The three needs are reflected through the self-declared motivations of participating in crowdsourcing translation in the *Yeeyan* community. Respondents to my questionnaire survey also acknowledged the contextual factors of the crowdsourced environment which can influence their participation (details see sections 4.2 and 4.4 in Chapter 4).

SDT can be considered as an extension of CET (Deci and Ryan, 2000). It does not solely focus on intrinsic motivation, but also other recognised motivational types including extrinsic motivation. SDT provides a taxonomy of motivational types based on the level of self-determination. Each motivation type has specific consequences for behaviour, performance and well-being. According to SDT, intrinsic motivation is truly self-determined, but extrinsic motivation is to a less extent self-determined, and 'amotivation' is fully non-self-determined (ibid). *Amotivation* is a state of lacking the intention to act and results from not valuing an activity, not feeling competent to do it or not expecting it to result in a desired outcome. The key point in SDT is that it provides a more comprehensive understanding of motivations. It not only analyses what pro-motivates people's activity, but also the de-motivations, which is important in organising crowdsourcing activities. It provides

guidance regarding which contextual factors (e.g. crowdsourcing process, evaluation metrics, and incentive mechanism) should be avoided or improved.

Moreover, SDT places intrinsic and extrinsic motivations on a continuum of autonomy where intrinsic motivation is expected to lead to higher levels of activity and higher quality of behaviour than all types of extrinsic motivation. However, the findings of my research are in contrast with the relation between motivations and quantity of contribution proposed in SDT. The self-declared data from the questionnaire survey demonstrates that extrinsic motivations lead to more time investment in participating crowdsourcing translation (see section 4.3 in Chapter 4).

GIT also argues that contextual factors influence intrinsic motivation and behaviour, but it focuses on only one contextual factor that influences the three psychological needs in CET: the *relevance of the task*, referring to the level that the task content or task context helps satisfy needs, wants and desires (Eisenberger et al., 1999).

CET/SDT scholars argue that financial rewards always have an undermining effect on intrinsic motivation and behaviour; on the other hand, GIT scholars argue that financial rewards can also have enhancing effects. In this research, the three theories are used to help analyse the effect of contextual factors on intrinsic and extrinsic motivations and on the participant's behaviour. Finding out what effect the contextual factors may have on motivation can inform the design and organisation of efficient future crowdsourcing projects.

2.3.2 Motivation studies in crowdsourcing translation

A limited number of studies have been conducted investigating the motivation in participating crowdsourcing translation. Conceptual frameworks are mostly based on the individual's internal psychological needs and external factors. Research methods include questionnaire surveys and documentary analysis.

Table 1 below shows the motivation categories and research methods used in prior studies.

Table 1. Motivation researches on crowdsourcing translation

Initiatives	**Conceptual Insights** **(Motivation Categories)**	**Research Method**
Rosetta Foundation (O'Brien and Schäler, 2010)	*Social motivations* (to support the initiative's cause; to support translation of information into a lesser-used language). *Personal motivations* (to improve language skills; to improve translation skills; to gain professional translation skills). *Cognitive surplus* (to gain intellectual stimulation).	Questionnaire Survey (Likert-scale rating)
Free Open Source Software (Lakhani and Wolf, 2005)	*Enjoyment-based intrinsic* motivation (intellectually stimulating) *Obligation/community-based intrinsic motivations* (believe that source code should be open; give back to the community; enjoy working with the development team; enhance reputation in the community; beat proprietary software). *Extrinsic-based motivations* (improve skills; user/personal needs; enhance professional status)	Questionnaire Survey
Wikipedia (Dolmaya, 2012)	*Intrinsic motivations* (1. *Political and social needs*: to make content available in another language; to bridge knowledge gap. 2. *Community motivations*: help support the organisation; community identification); *Extrinsic motivations* (to attract translation clients; to enhance reputation as translators; to gain more translation experience)	Questionnaire Survey

TED (Olohan, 2014)	*Pure altruism* (a desire to increase the supply of global knowledge; effecting social change). *Impure altruism* (the feel-good factor or the sense of satisfaction derived from altruistic behaviour). *Community identification.* *Personal motivation* (to enhance learning). *Enjoyment.*	Documentary Analysis (qualitative analysis of published translator's discourse, i.e. blog entries)
Scientific Memoirs[9] (Olohan, 2012)	*Pure altruism* (increase the public good of scientific knowledge) *Impure altruism* (sense of satisfaction; enhanced self-esteem or social esteem; to support non-for-profit action; community identification; to gain a degree of influence over the organisation) *Reputational concerns*	Documentary Analysis (fragmentary evidence drawn from letter exchanges between the publisher and contributors, and the *Preface* of the journal)

Drawing on previous studies, my research aims to build a more complete picture of motivational factors in crowdsourcing translation. Consequently, a *questionnaire survey* is used to deduce the motivating and de-motivating factors that influences people's participation in crowdsourcing translation projects in *Yeeyan*.

[9] Scientific Memoirs is a journal published in the 19th century for the advancement of British science. It was devoted to the publication of scientific translations and the international exchange of scientific ideas. Contributors (those who had been recognised) consisted of employees of the publishing company, scholars in science domain, and female translators (because women in the early 19th century had fewer opportunities for participating in scientific activities, they were more engaged in acquiring the linguistic proficiency that their male counterparts often lacked).

2.4 Project management in Translation Studies

Section 2.1.2.2 presented crowdsourcing as a whole from an organisational perspective. However, translation crowdsourcing includes more factors such as *workflow* and *communication* features, both of which are topics of my research.

2.4.1 Workflow of collaborative translation projects

TEP (translation – editing – proofreading) is one of the established models in the Language Services Industry (Kelly et al., 2011). It is also known as the 'waterfall' approach (Dunne, 2011) in which each step in the workflow waits for a hand-off before moving to the next one. The editing and proofreading of a project does not start until the translation is finished.

This model may impair translation quality at times, particularly in cases when speed is prioritised (Drugan, 2013:31). Neither will it fit in the localisation of online content, e.g. 'websites are dynamic in nature, with constant addition, changes and updates to source texts that require timely and efficient localisation workflows' (Jiménez-Crespo, 2017:63). However, as reflected in quality standards of translation services, a review (in EN 15038) and revision (in ISO 17100) stage is compulsory in the production process. Hence, to adapt to the broader context of translation services in the globalised environment (i.e. text type, translation purpose, and the prioritised requirement of the service – time/cost/quality), a flexible dynamic workflow that captures both translation and review/revision elements is needed.

DePalma and Kelly (2011:381) proposed a community approach by including a range of agents in the workflow (Figure 6): 'With collaboration, translation itself and revisions of translated work proceed in parallel with source content creation and editing, greatly reducing the lag between development and authoring, on the one hand, and completion of a localised product and publication of translated content, on the other hand'. The new model enhances the potential for collaboration by increasing value and reducing the

turnaround time.

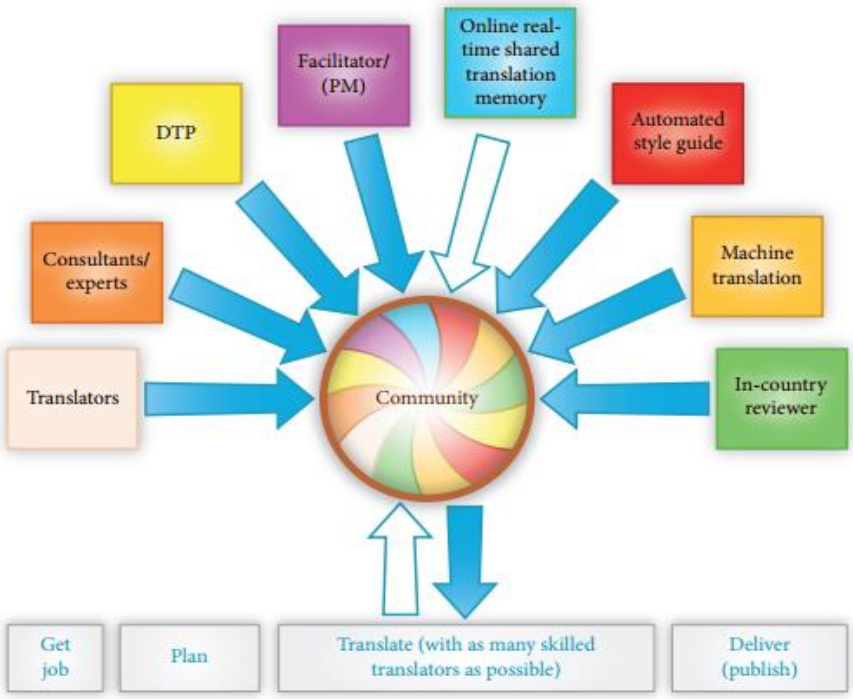

Figure 6. Localisation/translation workflow for collaborative translation in crowdsourced environment in comparison with TEP workflow (DePalma and Kelly, 2011:380)

Another change in the community approach is that it has re-allocated the responsibility for the successful completion of a translation process. Traditionally, the success of a translation project lies on the shoulders of professional translators and revisers in collaboration, yet the new model reallocates the responsibility by assigning some weightings to 'developers and managers who need to prepare successful workflows, management procedures, community building and the different levels of authority and participation' (Jiménez-Crespo, 2017:63-64). Apart from the weights carried by a range of agents involved in the *production* stage, the new paradigm highlights the *preparation* stage in terms of project plan. It refers to all the operations carried out prior to the translation production stage, which includes the involvement of managers and/or the agents who are responsible for the preparation work (i.e. analysis, documentation, terminology, resources

and tools, etc.) (Gouadec, 2007). So far, few empirical studies have examined the influence of workflow and management on the success of the project. In this PhD project, I focus on the workflow and management procedures in collaborative translation projects in the *Yeeyan* community with specific focus on their influence on the quality of translation (see Chapter 5 and 6.4).

'Content selection' and 'the selection and development of platform' are found to be essential in the *preparation* stage in the crowdsourced environment. Types of content that are highly amenable to crowdsourcing are those with embedded communities (e.g. software and social networking sites), which are familiar to large groups of people, and with cultural values (Morera-Mesa, 2014:136). In terms of the selection and development of a platform, Dombek (2014) found that the shortcomings of crowdsourcing platforms can be detrimental to the engagement and motivation of volunteers. Empirical research is needed to explore the workflow to find out more procedures that contribute to the success of crowdsourcing translation.

DePalma and Kelly (2011:381) indicated that the new community workflow has already been widely used by both suppliers and buyers of language services to 'execute large projects in short bursts or as part of systematic translation and localisation efforts'. However, the new model also poses a number of challenges, especially in the crowdsourced environment. With more agents sharing the responsibility of collaborative projects, the significant trust and the power-relationship between agents appears to be one of the most unstable elements in crowdsourcing (Jiménez-Crespo, 2017). Moreover, given that crowdsourcing involves an overhaul of the traditional sequential translation process, factors such as provenance, ability and experience of participants, the use of technology, and project management approaches, all need to be carefully considered.

Jiménez-Crespo (2017:132) conducted a comprehensive review of the existing crowdsourcing models, as well as industry and scholarly publications, and identified a list of practices that help guarantee quality in the crowdsourced environment:

1. *Translation – Editing – Proofreading (TEP) approach*. It implies a hierarchical pattern by involving professional or good performers to the revision or management, e.g. *TED Open Translation* (Cámara, 2015), or *The Rosetta Foundation* (O'Brien and Schäler, 2010).
2. *Preselection with exam*. It is often found in non-profit and paid crowdsourcing initiatives, such as *Kiva*, and *MOOC*[10] (Cao, 2015).
3. *Selection of the participants in the crowd*. Open and closed models of crowdsourcing represent different approaches to this issue, but even in open models sometimes filters are set in place. This approach usually specifies the desired profile of the participants, such as knowledge in a specific area, experience in translation or experience in consuming the source content or products.
4. *Continuing evaluation of participant performance*. Participant performance can be evaluated either manually or automatically. The evaluation is related to the upgrade/development of participants in the community.
5. *Creation of cloud workflow solutions with built-in measures*. It refers to using customised platforms to manage the translations under different contexts, such as MNH-TT (Babych et al., 2012).
6. *Establishing common resources such as glossaries, guidelines, norms or discussion forums*. It can resemble the professional practices in workflow management solutions (Morera-Mesa, 2014; Orrego-Carmona, 2012).
7. *Combination of automatic quality checking with professional human translator*. Automatic evaluation of MT methods can be usually used (Zaidan and Callison-Burch, 2011); and professional translators contribute to select the best rendition.
8. *Presence of language experts or community managers*. Language experts and community managers serve to oversee the crowdsourcing process and the overall quality of translation (DePalma and Kelly, 2011); for example, the *Microsoft Collaborative Translation Framework* (Aikawa

[10] MOOCs (Massive Open Online Courses) are free online courses available for anyone to enrol on. A course on Open Translation Tools and Practices (OT12) was launched in 2012, in which participants explored a range of online open translation tools and collaborated in the translation and subtitling of open education resources.

et al., 2012).

9. *Quality loops after the release of the text*. This approach refers to the participatory mechanisms in which translation audiences can edit or report translation issues, such as *Wikipedia* and the discussion boards in FOSS localisation communities.
10. *Iterative/redundancy voting*. *Facebook* is the typical example adopting this approach by asking end-users to vote for the best/better translation. However, the approach involves segmentation of the source text and it is not suitable for longer content, such as paragraphs or entire texts (Jiménez-Crespo, 2011).
11. *'Many eyes' principle*. It starts from the FOSS communities: with a large group of beta-testers and co-developers, problems can be easily spotted and fixed (Raymond, 2001). It is available only in totally open systems in which the function – 'freezing' of the translation is not incorporated, similar to the one in FOSS.

These approaches vary to a great extent, but all of them share the need to develop alternative mechanisms under different scenarios. Some of them are merely extracted from existing experiences without empirical examination of the translation quality resulting from these approaches. Nevertheless, they provide suggestions for the future organisation of crowdsourcing translations and inspire new perspectives in researching crowdsourcing workflow and translation outputs. In this PhD project, I also attempt to find out the feasibility of some of the proposed approaches in the *Yeeyan* community and their influence on translation quality with empirical data.

2.4.2 Communication management in collaborative projects

Scholars have recognised the importance of communication management in translation and localisation projects (Dunne, 2011; Matis, 2014). It is generally believed that PMs spend 80% of their time communicating (Treasury Board of Canada Secretariat in Dunne, 2011:189). Furthermore, it is also generally accepted that efficient communication can greatly reduce the workload of the PM. More importantly, it reduces the risk of project failure (Matis, 2014). To be more specific, Matis (2014:169) highlighted that

requesting all actors involved in the project to confirm the receipt of e-mails and files, as well as the acceptance of the delivery date can avoid the problem of certain tasks not having been processed by the designated people at the right time. Also, constant communication ensures that team members work in accordance with the project brief and that the project progresses according to the initial plan. It is important that such communication is not limited to the beginning of the project – 'regular monitoring should be undertaken to ensure the project progress' (ibid:170).

Tsvetkov and Tsvetkov (2011) indicated that, for PMs in localisation projects, especially international-wide multilingual projects, two factors should be considered: individual personality, and cultural background of peers. As these two factors influence people's 'way of thinking, […] decision-making process, the value they ascribe to individual goals versus group goals, and their perception of time' (ibid:190), identifying the characteristics of the two factors will improve communication. Moreover, a communication plan should be established to specify the information needs for stakeholders: when, how often and in what formats they need this communication.

Two types of meetings were seen to work well for translation and localisation projects: a quick ten-minute check-in meeting with the project team each morning (in-house meeting, telephone, or video meeting according to the geographical setting of the project members); and a weekly meeting which covers greater detail of the project updates (Lencioni in Dunne and Dunne, 2011:206). The weekly meeting usually has a strict agenda, including reporting old tasks, introducing new tasks, and a lesson-learnt session.

As can be seen from the discussions above, the communication carried by the PM within a translation/localisation project is very important. However, the role of other actors' performance in communication during the project is very often overlooked.

Soller (2001) developed a series of task-oriented dialogue acts to capture students' interaction in the collaborative learning environment. It defines acts regarding 'request, inform, motivate' as signs for *active learning*; 'task, maintenance, acknowledge' as *coordinative conversation*; 'argue, mediate'

as signs for *creative conflict* (for more explanations, please see Figure 7). The study showed that communication enhances the learning experience because of the knowledge exchange hidden in the dialogue acts. It offers insights into a new research method which focuses on investigating collaborative groups. On the one hand, the analysis of the interaction helps draw inferences of collaborative groups – their strengths and weaknesses – by investigating how many or what kinds of questions are asked, as well as how many answers of which level of quality are provided by individuals or the group (ibid). The result can be used to determine the appropriate method and strategies to further group work. On the other hand, the method includes all actors in the collaborative environment into the picture of communication management.

Definitions of Collaborative Learning Conversation Skills and Subskills

Active Learning	**Request**	**:** Ask for help/advice in solving the problem, or in understanding a team-mates comment.
	Inform	**:** Direct or advance the conversation by providing information or advice.
	Motivate	**:** Provide positive feedback and reinforcement.
Conversation	**Task Maintenance**	**:** Shift the current focus of the group to a new subtask or tool.
	Acknowledge	**:** Support group cohesion and peer involvement.
		: Inform peers that you read and/or appreciate their comments. Answer yes/no questions.
Creative Conflict	**Argue**	**:** Reason (positively or negatively) about comments or suggestions made by team members.
	Mediate	**:** Recommend an instructor intervene to answer a question.

Figure 7. Dialogue acts in collaborative learning conversations (Soller, 2001:6).

Babych et al. (2012) emphasised the role of communication in scaffolding the trainee's experience. It is relatively rare in Translation Studies, as communication so far is only introduced from the professional project management's perspective. The researchers reflected on collaborative activities on the translation platform – *Minna no Hon'yaku* (Translation of/by/for All). They incorporated communicative functions into the existing platform, along with other functions, including organisational function, motivational function and resource function, hereby extending the platform to a system that was suitable for translator training (*MNH-TT*).

The core roles in *MNH-TT* were identified as: researcher, terminologist, translator, reviser, reviewer, and proofreader. Based on Soller (2001) and other previous social interaction researches, in *MNH-TT*, dialogue acts were defined and grouped thematically by Babych et al. (2012:9) as follows:

- *Information-acts* bear on the provision and discussion of information related to the substance of the translation;
- *Maintenance-acts* maintain the flow of dialogues and promote the cohesion within the team;
- *Status-acts* bear on the progress of the work schedule;
- *Role-acts* relate to the 'staffing' of the translation team.

Acts are represented as a menu in *MNH-TT*. Each time a message is posted, the sender chooses the appropriate act type to classify the contribution. The interaction between participants during the translation can be analysed in correlation with significant modifications to the translation product.

Inspired by previous studies, this research focuses on the communication between participants in the crowdsourced environment and uses a modified model of dialogue acts in *MNH-TT* to capture the participants' behaviours (for more details, see Table 4, in Methodology Chapter). The new model combines both aforementioned *promoting* and *monitoring* dialogue acts from the perspective of project management, as well as the *social* interactions and *knowledge exchanges* between **all actors** involved in the project.

2.5 Quality Assessment in Translation Studies

2.5.1 Translation Quality

2.5.1.1 The notion of translation quality in academia

The definition of quality is highly elusive, especially for language-based professions, as it may be perceived very differently according to one's position or role (Fox, 2009). For centuries, translation quality has been debated in a variety of contexts (Brunette, 2000:169), yet theorists and professionals alike still agree that there is no objective quality measurement

(Drugan, 2013:35). The question of ‘who is to judge the quality of the translation’ plays an important role in quality assessment by service providers, client organisations, or the end user, who may each hold different quality criteria. It also depends on individual tastes and preferences. Hence, as Fox (2009:24) noticed: ‘the quest for quality will have to integrate all these different facets in varying degree.’

Quality is not a static notion. It constantly evolves and is often framed by practices and discourse in real-world scenarios (Jiménez-Crespo, 2017). As Jiménez-Crespo (2013a:103) wrote, quality evaluation tasks are ‘built up internalised frameworks of quality that guide subjects’ decisions, even when they might lack operative theoretical foundations’.

The alternative translation production models, such as machine translation and crowdsourcing translation, challenge the criteria for translation quality – i.e. whether top, publishable quality is ideal for all scenarios (Jiménez-Crespo, 2017) with localisation companies frequently using tiered pricing structures paired with various translation technology solutions and quality levels (VideoLocalise, 2017). The question of what constitutes ‘acceptable’ quality levels rather than how to achieve the highest or best quality possible has been a focus in the translation industry of late. Concepts such as ‘fit for purpose’ translation, or ‘good enough translation’ are frequently discussed and have already been supported by leading figures in the industry (Prioux and Michel, 2007).

In the earlier stages of crowdsourcing translation, professionals would frequently complain that the highly aspirational notion of ‘top quality’ is an unachievable and imaginary myth (Jiménez-Crespo, 2017). The main dispute was the inability to produce ‘sufficient’ quality if professionals were not involved (Kelly, 2009), often pointing at the prevalent number of errors in non-professional translations (Bogucki, 2009; Khalaf et al., 2015). With the expansion of the emerging alternative translation models, such as volunteer or paid crowdsourcing by either professionals or bilinguals, as well as volunteer or professional post-editing of machine translation, quality started to be conceptualised in a continuum, ‘with the enduring and critical at one

extreme, and the ephemeral and inconsequential at the other', in which traditional translation and translator education approaches the former and struggles to keep pace with the 'fast and fit' imperative of the latter (Garcia, 2014:430).

2.5.1.2 Translation quality standards in the translation service industry

In terms of the quality standards in the professional industry, as Görög (2017) indicated, different translation service providers use different metrics, which makes it difficult to compare vendors, translators, projects and to benchmark translation quality with industry averages. Drugan (2013:74-75) summarised a series of standards recognised by language service providers that she interviewed, including ISO 9000 series, DIN 2345, EN 15038:2006, ASTM F2575-06, and GB/T 19363. However, because some of the standards are outdated and replaced by the new ones, I will briefly introduce the **currently most used standards** in the industry.

EN-15038:2006: European Quality Standard for Translation Service Providers: the most outstanding feature in the standard is its definition of translation process **where quality is guaranteed not by the translation which is just one phase in the process, but by the fact of the translation being reviewed by a person other than the translator**. It also specifies **professional competence of each of the participants** in the translation process, mainly translators, revisers, reviewers and proofreaders. (European Committee for Standards, 2006)

ISO 17100:2015: ISO 17100 superseded EN 15038 in 2015. Certification and ongoing auditing are required. The new standard specifies requirements for all aspects of the translation process directly affecting the quality and delivery of translation services:

- **Resources**: requirements for the qualifications and professional competences of translators, revisers (bilingual editors), reviewers (monolingual editors), project managers, etc.;
- **Pre-production process and activities**: requirements for quotation, enquiry, feasibility assessment, client-company

agreement, administrative activities, technical aspect of project preparation, linguistic specification, etc.);

- **Production process**: requirements for the different stages of the production process, including project management, translation, revision, review (monolingual editing by domain specialist, and proofreading (target language content revision and correction before printing), and the final verification and release of the translation by a qualified project manager. It highlights that any translation service under ISO 17100, must include, as a minimum, **translation and revision (bilingual editing)**);
- **Post-production processes**. Different from EN 150382, it highlights the importance of interacting with the client, both in the initial translation services agreement and in managing possible modifications, claims, feedback, customer satisfaction assessment and closing administration). (International Organisation for Standardisation, 2015)

ASTM (American Society for Testing and Materials) F2575-14: an updated version of F2575-06, it is a standard guide for translation quality assurance. The standard does not provide specific criteria for translation or project quality, as these requirements maybe highly individual. Instead, it specifies **parameters that should be considered before translation begins**, and informs stakeholders about the basic quality requirements are in need of compliance. (American Society for Testing and Materials, 2014)

National Standard of the People's Republic of China GB/T 19682-2005, GB/T19363.1-2008: translation-specific requirements which was developed with reference to the EN standard. It provides a comprehensive range of **quality features**, which is elaborated in section 3.2.4.3 in Methodology Chapter. The standard also includes a detailed section on linguistic quality elements (e.g. how foreign names, or units of measurement must be translated into Chinese) (State Administration for Quality Supervision and Inspection, 2005, 2008).

What is very important to appreciate – which is in fact an aspect often missed

by freelance translators debating the relevance of such standards – is that the industry standards are mostly process-oriented, focusing on establishing and maintaining a process of translation, review, and approval. These, when followed by qualified professionals, will consistently result in translations that meet both customer expectations and industry requirements (Morningside Translations, 2016). They demonstrate the increased attention to quality at different stages of the translation process, which further emphasises 'one of the main differences between theoretical and real-world approaches' (Drugan, 2013:74). The following sections discuss both academic and professional approaches to translation quality assessment.

2.5.2 Translation Quality Assessment (TQA)

2.5.2.1 Academic approaches to translation evaluation

Translation evaluation was merely a simple commentary of literary works until the second half of the 20th century (Secară, 2005). Due to the lack of explicit criteria and sound methodology, the evaluation was mostly conducted subjectively and consisted of discussions dealing with little more than the translations' faithfulness to the original (ibid). The situation began to change in the 1970s with the growing importance of Translation Studies as a discipline and the theoretical development of translation evaluation. Eugene Nida (1964) was one of the first scholars who discussed translation quality and proposed to use the final translation product for evaluation. He prioritised the readers' reactions to the translation when evaluating this product and indicated that 'good' or 'bad' translation can be identified according to its capacity to make the translation's readers react in the same way the source audiences did. The theoretical framework of such evaluation – 'equivalent effect' or 'dynamic equivalence' – is still debated in today's Translation Studies (Secară, 2005). Vermeer's (2004) 'Skopos' theory is a new development of translation evaluation in the 80s. In his theory, the purpose of translation was the prime consideration when evaluating it. The functionalist approaches in the 90s were more widely used, which emphasised the target reader/end-user's position in evaluating a translation – whether it fulfils its function in the target text (Schäffner, 1998).

The theories mentioned above laid the foundation of translation quality which facilitates the establishment of an extensively recognised evaluation system. House (2001:200) called for ‘intersubjectively verifiable evaluative criteria’ which needs to be achieved through extensive empirical studies analysing large multilingual corpora of translated texts. Academics put forward TQA approaches in various ways. However, few of them published ‘detailed, reproducible TQA models for human translation with an indication of the text types on which they were tested’ (Drugan, 2013:46). The following section reviews a few specific models in the academic field.

Larose (1987) put forward a teleological model for translation assessment, in which they emphasised the importance of objective or purpose of the translation in assessing quality: how far the translator’s purpose matches the original author’s intention, rather than the preferences of the evaluator. A hierarchical structure is proposed:

- **microstructural** – the lowest level of assessment, which concerns ‘forms of expression’ at the sentence or sub-sentence level, such as graphic, syntactic and lexical factors;
- **macrostructural** – assesses beyond the sentence level, which concerns the semantic structure of discourse content, or cohesion;
- **superstructura**l – the highest level, which assesses the overall structure of discourse, including the narrative and argumentation organisation, text type, and function.

Extra-textual factors are also valued in the model, such as the translation’s aim, client brief, and other requirements the translator needs to fulfil. Furthermore, Larose’s hierarchy model is used to weigh up the errors in translation: according to the functional or structural importance of the error and the level at which it occurs, microstructural errors are weighted as less significant than errors in macro and superstructural levels.

House (2001) proposed three approaches for TQA: anecdotal and subjective, response-oriented, and text-based.

- The first approach, **anecdotal and subjective**, applies to the early stage of Translation Studies when translation is devised by ‘practising

translators, philosophers, writers and many others'. It is atheoretical, and as House (ibid) stated, the approach normally rejects the establishment of general principles for translation quality.

- The second approach, **response-oriented** model, is based on the equivalence and equivalent effect theories, in which the measurement focuses on the response of target readers/end-users of translations and compares it with the response of source text readers/users –for example, in the study of examining people's acceptance of fan subtitling (Orrego-Carmona, 2015).
- The third approach, **text-oriented** model, relates to Reiss's (1989) *text-type theory*, Vermeer's (2004) *skopos theory*, and Nord's (1997) *text analysis model*. This approach focuses on a detailed analysis and comparison of source texts and target texts, or the target texts alone.

Al-Qinai (2000) suggested an 'eclectic' model of TQA which is based on the textual analysis of a series of parameters:

- **textual typology (province) and tenor** – it refers to the linguistic and narrative structure of the source texts and target texts, and the textual function;
- **formal correspondence** – it deals with the presentation of the translation compared with the source text (e.g. paragraph division, punctuation);
- **coherence of thematic structure** – it examines whether the thematic development of the target text is consistent with the source text;
- **cohesion** – it examines whether the translation follows the rhetoric and sequence of thought in the source text;
- **text-pragmatic (dynamic) equivalence** – it looks at the equivalent effect between the target reader and source reader;
- **lexical properties** – it caters for the register and linguistic features (e.g. specialised terminology, idioms, and collocations) through the comparison of source text and target text;
- **grammatical/syntactic equivalence** – it is examined through the comparison of source text and target text with regard to word order and agreement (e.g. number, gender, person).

Al-Qinai's model evaluates translation as a product rather than as a process. The parameters in the model imply the empirical and objective side of evaluation. However, it is difficult to apply in the real world because it mostly depends on the comparison of source text and target text and certain information about the source text (e.g. production of the original text) and the translation process (e.g. working conditions, client brief) is notoriously difficult to obtain. Moreover, it takes huge effort and time to assess translation using the model. For instance, Al-Qinai used 16 pages to assess a translation of 300 source text words, which makes it impractical for real-world applications (Drugan, 2013).

Williams (2004) developed a framework of TQA based on discourse analysis. It involves the analysis of the source text and target text individually to establish the 'argument schema, arrangement and organisational relations' of each, the comparison of source text and target text to assess the 'transfer of argument', and the overall quality statement. The model is based on the error analysis by addressing the acceptability threshold of errors under different standards (i.e. the level of errors which can be tolerated in a translation). Four standards of translation were proposed:

- **publication standard** – the translation accurately renders all components of the 'argument schema' and meets the conventions of target language, domain and user-specific parameters, and contains no critical or major errors;
- **information standard** – the translation renders all the 'argument schema', and meets the conventions of target language, domain and user-specific parameters, and contains no critical errors;
- **minimum standard** – the translation transfers all the 'argument schema' and contains no critical errors;
- **substandard** – the translation fails to convey the 'argument schema', contains at least one critical error, and/or does not meet the requirements of domain and user-specific parameters.

The definition of an error in Williams' model is different from the standard industry quality control which differentiates between critical, major and minor

errors. It relates to the ‘usability of the translation [where an error] impairs the central reasoning of the text’ (ibid:133), rather than the accuracy or fidelity of the translation. The model is insightful in quantifying translation quality, but a danger of subjectivity is embedded in the evaluation as it largely depends on the human understanding of the ‘argument schema’.

O’Brien (2012) proposed some basic parameters in dynamic quality evaluation: ‘content type, communicative function, end user requirements, context, perishability or mode of translation generation’. When benchmarking the method for quality evaluation, special attention is paid to the variability of the communicative function of the translated text and content type. Three communication channels of translated content were suggested:

- **B2C** (Business to Consumer),
- **B2B** (Business to Business),
- **C2C** (Consumer to Consumer).

This distinction matters because the nature of the communication channel will impact on translation quality expectations, and more importantly, it helps select the quality evaluation model. The C2C model caters for multilingual user-generated content, which is consumed by multilingual users. Internal communication, for example in the B2B and C2C models, will require lower quality than external communication model – B2C.

Content type is profiled according to three factors: **utility**, **time** and **sentiment**. *Utility* refers to the relative importance of the functionality of the translated content; *time* refers to the speed with which the translation is required; *sentiment* refers to the importance of impact on brand image. Translation content can be profiled according to the importance of these three factors and consequently, together with the communicative function of the translated text, the quality evaluation model can be decided: whether to use the conventional adequacy/fluency evaluation, community-based evaluation, or error-based evaluation. O’Brien’s model is superior to previous text-based and functional approaches to evaluation because it takes more contextual factors into consideration, which responds to the new forms of translation in the globalised multilingual environment.

2.5.2.2 Professional approaches to translation evaluation

Academic models of translation evaluation reflect the development of Translation Studies in both theory and practice. The professional approaches aim at a more straightforward and more systematic type of translation evaluation, focusing on the 'creation of explicit and applicable correction scales' to fulfil 'the commitment to deliver error-free translations to clients' (Secară, 2005:39). Thus, a series of error-based TQA models were developed.

The *Canadian Language Quality Measurement System* (Sical) is an early innovation in categorising errors for TQA. It was developed by the Canadian government's Translation Bureau for the assessment of instrumental translation. Its successive versions incorporated a scheme based on the quantification of errors on a twofold distinction between (1) translation and language errors and (2) major and minor errors (Williams, 2004:3). A *major error* in this model was defined as follows:

- Translation: complete failure to render the meaning of a word or passage that contains an essential element of the message; also, mistranslation resulting in a contradiction of or significant departure from the meaning of an essential element of the message.
- Language: incomprehensible, grossly incorrect language or rudimentary error in an essential element of the message.

A threshold based on the number and type of errors in a 400-word passage is used to define the translation acceptability: 'A – superior (0 major errors/maximum of 6 minor); B – fully acceptable (0/12); C – revisable (1/18); and D – unacceptable' (Williams, 2004:3). The standard shows that Sical uses sample analysis at the word and sentence level, not on the overall text. It received criticism regarding the imprecision of the specific number and type of errors in not treating the translation as a whole (Williams, 2001). In addition, it contains 675 error types (300 lexical and 375 syntactic) (Melis and Hurtado, 2001:274), which makes it difficult to apply (Secară, 2005).

SAE J2450 was developed in the 90's by SAE (Society of Automatic

Engineers) in collaboration with GM (General Motors) to provide an objective assessment of linguistic quality for automotive service information regardless of language or process. Based on seven error categories, it focuses more on content problems that may influence the understanding of the message, than on style problems. The error categories are:

- wrong term
- syntactic error
- omission
- word structure or agreement error
- misspelling
- punctuation error
- miscellaneous error

Each category has a certain weight that decides the seriousness of the error. 'The final score is obtained by adding up the scores of the errors and dividing the result by the number of words in the text' (Secară, 2005:40). The model is questioned because it mainly caters for terminological evaluation and lacks a defined threshold of the acceptability of translation (ibid).

BlackJack is a translation evaluation tool developed by the British translation agency ITR. According to Eckersley (2002), it aims to reduce the effort and time spent by reviewers/evaluators, and to provide a formative evaluation through a standard checklist. 21 error categories are included in BlackJack covering both content and style problems. Weight is added to each error category in the form of a numerical value to represent the seriousness of the error. The tool automatically assigns the weight to each error identified by the human evaluator, and then computes a final score of the translation quality at the end of evaluation. The tool is an advanced step of *SAE J2450* with the enriched error categories (from 7 general categories to 21 more specific categories) (Secară, 2005). However, some of the error categories are unjustifiable. For example, categories such as 'spelling error' and 'spelling deviation' are treated separately in the model with the same numerical value. This is confusing because they should be considered as different errors only if they have different values (ibid:41).

The *MeLLANGE* (**M**ultilingual **eL**earning in **LANG**uage **E**ngineering) project is a research project that brought together academic and industry partners across Europe, with the aim to 'adapt vocational training for translators and other language professionals to meet the new needs that flow from globalisation and increased cross-cultural communication: an increased need for the management of intercultural language resources' (MeLLANGE, ND)[11]. One of research outcomes is the *MeLLANGE* error typology that aims at 'helping trainers on how to comment the translations performed by their students' (Secară, 2005:43). The error typology is a hierarchical scheme that covers content, language, terminology and lexis, hygiene (e.g. spelling and punctuation), register and style issues. The main categories are further divided into 30 subcategories (Figure 8). Different from the error typologies used in the professional industry, *MeLLANGE* does not rate the translation quality, but instead seeks to identify and classify types of errors to provide objective and formative feedback in translator training. Hence, it attempts to cover translation issues in as much detail as possible. However, since it is a project largely based on Romance languages, some of the subcategories may not suit other language systems. For example, the subcategory *Inflection* and *Agreement* under *Language* category is superfluous in Chinese due to the absence of gender and number in this language. Therefore, if the *MeLLANGE* error typology is to be used across the global translation institutions, it would need to be further adapted according to the language feature and real scenarios.

[11] The MeLLANGE Project is available from http://mellange.eila.jussieu.fr/.

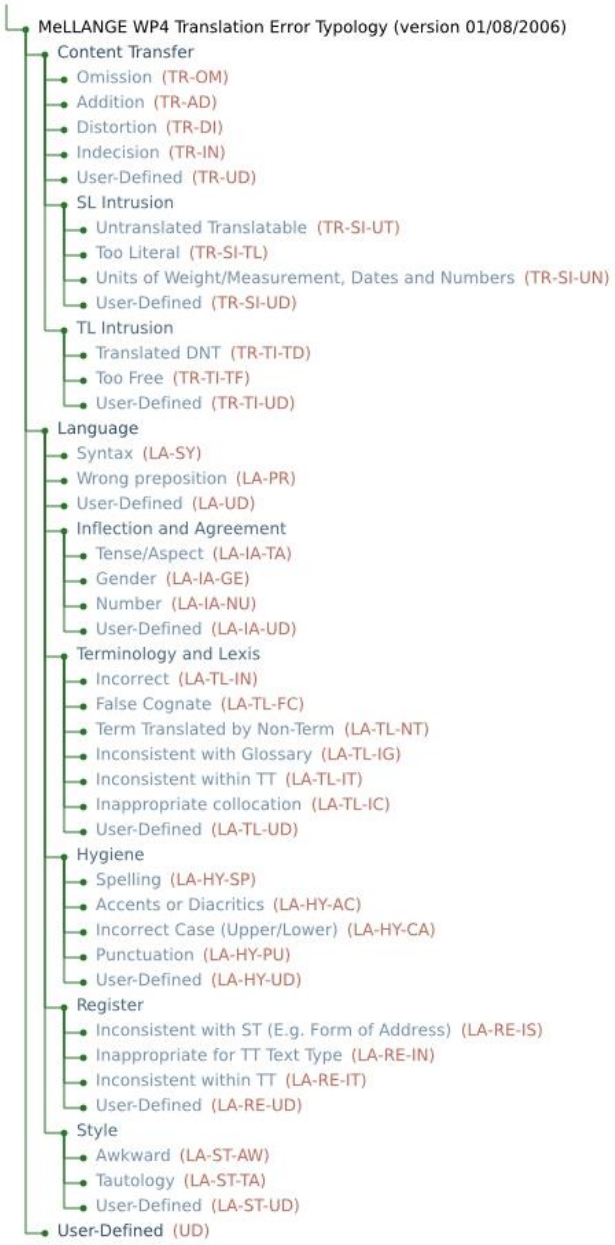

Figure 8. *MeLLANGE* error typology[12]

Recent developments in translation evaluation tend to be dynamic and highly customisable to fit in the broader context of translation practices. The *Multidimensional Quality Metrics* (MQM) system was developed in the European Union-funded *QTLaunchPad* project to address a full range of issue types. Figure 9 presents the 'core' error categories in MQM. As can be

[12] MeLLANGE error typology (available from http://mellange.eila.univ-paris-diderot.fr/public_doc.en.shtml).

seen, the error categories are arranged hierarchically. A set of threshold and severity levels of errors are included in the model. The ‘core’ consists of 20 error types and it is extensible when needed. According to Lommel et al. (2014:458), ‘Every node (including very high-level ones like Accuracy and Fluency) can serve as an issue type, and children of an issue represent specific cases of the parent issue. As a result, an MQM metric can be declared at various levels of granularity.’ The innovation brought by the MQM lies in its division of errors into both a general level and a detailed level, which is useful in diagnosing the problems that fit both theoretical considerations (higher branch in the tree graph) and specific needs for improvement (lower branch in the tree graph).

‘Accuracy’ represents the content issues, while ‘Fluency’ and ‘Style’ cater for language issues. ‘Verity’ is a novel contribution of MQM as a professional approach. It refers to ‘extra-linguistic truth correspondence in the texts (ibid:459), and requires the evaluator to consider the text in its intended/target audience environment. ‘Design’ refers to the issues related to the physical presentation of text, typically in a ‘rich text’ or ‘markup’ environment, which is of particular importance in localisation practices.

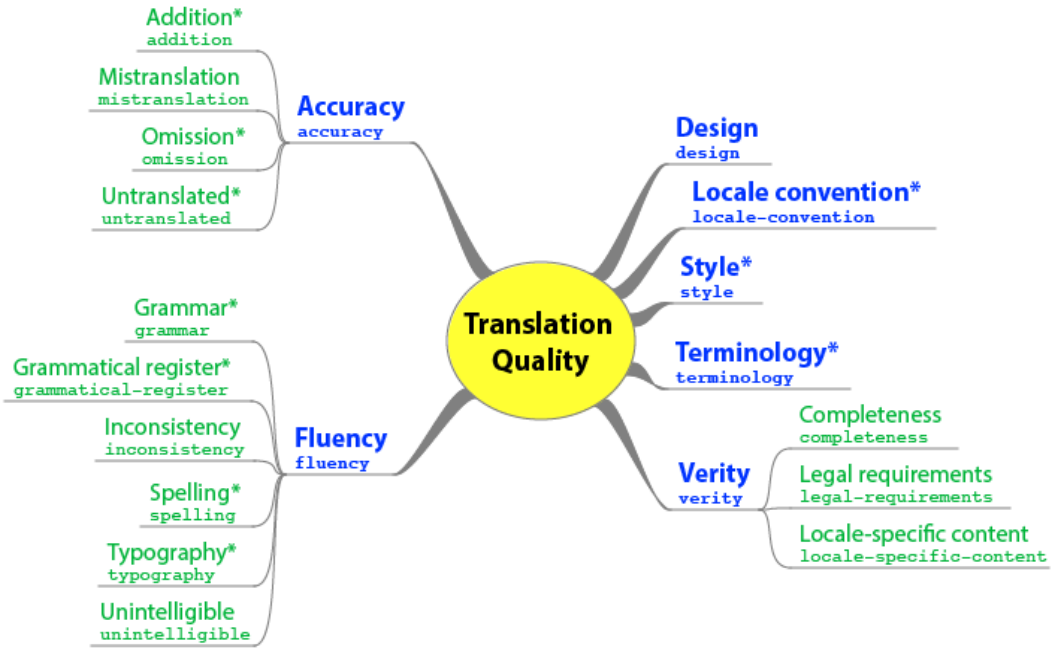

Figure 9. A tree graph of the hierarchy of the MQM core error categories in version 1.0[13]

The *Dynamic Quality Framework* (DQF) was developed by Translation Automation User Society (TAUS) approximately at the same time as MQM. Following O'Brien's (2012) work on the parameters of the dynamic quality evaluation, the DQF offers a more flexible approach than the previous static quality evaluations based on the three parameters of *Utility*, *Time* and *Sentiment* (UTS) (Görög, 2014). Prior to quality evaluation, the *TAUS Content Profiling* tool allows an evaluator to conduct a UTS rating of the source content and specify the communication channel of translation. The result is then mapped to suggest a specific quality evaluation model to suit the evaluator's needs. After content profiling, a quality evaluation project can be set up in the DQF tools. According to Görög (2014:449), 'One of the aims of DQF tools is to standardise the evaluation process and to make it more objective and transparent. The benchmarking and reporting functions provide users with a wealth of information on quality problems related to certain language pairs, text types, industries or domains.' The DQF enables

[13] A tree graph of the hierarchy of the MQM Core error categories in version 1.0 (available from http://www.qt21.eu/mqm-definition/definition-2015-12-30.html) (Lommel et al., 2014).

evaluators to assess quality from three different sources: *Adequacy*, *Fluency* and *Error Typology*. The *adequacy* and *fluency* tools rate the transfer of source content and the target grammar on a scale of 1-4, which are straightforward to use. In terms of the *error typology*, four main error categories are included in the DQF tool: Language, Terminology, Accuracy and Style. The errors can be annotated on four severity levels: critical, major, minor and neutral. Neutral is applied when a problem needs to be tagged, but is not the fault of the translator. The DQF tools can be used to evaluate both human and machine translation. This framework can also be applied to measure post-editing productivity and to score translated segments based on an error typology. The advantage of the DQF tool lies in its standardised way of categorising and counting translation errors using the commonly-used industry criteria (ibid:452). However, García (2014) pointed out that the error typology in the DQF tools need more development, and in March 2016, TAUS published a new version of the DQF tools which is the result of the harmonised DQF-MQM error typology model, combining the analytical method of DQF with the MQM hierarchy of translation quality issues.

The most recent error typology was proposed by Fujita et al. (2017), which aims at providing a consistent categorisation of 'issues' (i.e. translation errors and problems) in **learner translation** in the context of **English-Japanese** translation. The typology preserved the top-level error categories of the *MeLLANGE* typology and created subcategories to cater for the specificities in the Japanese language and the translation training environment. Moreover, they also developed a *decision tree* as a navigation tool to support human decision making because the typology is expected to be used by both the instructors for translation quality assessment and for the learners to understand the diagnoses. The typology was empirically examined, which confirmed the coverage of the defined 'issue' types and the acceptance by both the instructor and the translation learners. Several of the issue types in Fujita et al.'s typology are recognised in my study, as well. For example, during the quality assessment of crowdsourcing translation outputs, I also identified *collocation* as one of the subcategories in lexical issue, and *text-clumsy* (meaning that the 'lexical choice or phrasing is clumsy, tautologous,

or unnecessarily verbose' (ibid:60)) is similar to my categorisation of '*redundant expression*' which is found to be a common problem among the crowdsourcing translators (for more details, please see section 7.3.2). In addition to *MeLLANGE*, Fujita et al.'s typology is one of the few empirically-tested typologies that focuses on translation training.

This study uses an adapted version of the *MeLLANGE* error typology to assess translation quality due to the specificities of the crowdsourced environment and the language characteristic of Chinese. The reason was that the other error typologies in the existing error-based TQA models did not fit in the context of the research. For example, due to the absence of gender and number in Chinese, an error category of the type 'Language: Inflection and Agreement' in *MeLLANGE* was superfluous; the 'Design' category in MQM (Multidimensional Quality Metrics) was not relevant, either, because participants in the *Yeeyan* community only perform translation-related activities and they are not responsible for the physical presentation of the translated text. Table 6 in the Methodology Chapter presents the customised error typology in this study.

2.5.3 Empirical research on crowdsourcing translation quality

Though quality has been a focus in Translation Studies, particularly for the evaluation of machine translation, there are few empirical studies on quality evaluation of crowdsourcing translation. Jiménez-Crespo (2017:154-155) summarised the existing quantitative studies on crowdsourcing translation quality, mentioning Pérez and Carreira (2010), Deriemaeker (2014), and Jiménez-Crespo (2013a; 2015).

Pérez and Carreira (2010) investigated the quality of the Spanish translations on Facebook. An error-based evaluation was used to identify patterns of inadequacies in the translated interface of *Facebook*. The result recognised calques, lexical selection, coherence, spelling and formatting as the main issues. In this PhD project, I assess quality of crowdsourcing translation outputs with more explicit error annotation. Significant error types in different text domains are identified, which are further used to propose solutions to these issues (see sections 7.2.2 and 7.3.2 in 6.4). I also investigate the

influence of workflows on translation quality.

Deriemaeker (2014) conducted an experiment on crowdsourcing translation of tourist brochures using *Duolingo* (a free language-learning and crowdsourcing translation platform). The crowd's output was compared with professional translations to annotate acceptability and adequacy. The result demonstrated that the number of errors between professional translation and crowdsourcing translation is similar; however, the types of errors in the crowd's output are weighted more seriously than professionals'. The experiment suggested that crowdsourcing translation is not of 'low-quality and very much understandable to readers' (ibid:48), which makes it adequate for information purposes. The successful crowdsourcing applications such as *Facebook*, *Mozilla Firefox*, and *Global Voices*, have already demonstrated that translations produced by the crowd are well received by the audiences. We cannot ignore the fact that collective intelligence can produce translations almost as good as those of professional translators, which is difficult for some people to accept. My research is set to assess the quality of crowdsourcing translation in different text domains (business, health and literary) to provide empirical evidence for raising the recognition of the crowdsourcing model.

Jiménez-Crespo (2015) conducted an experiment to investigate the influence of translation methods on the translated products. The study recruited 75 translation students and divided them into three groups: the first group directly translated segments from *Facebook* (subjects were unaware that the segments were extracted from the *Facebook* interface); the second group selected translations from the potential range of terminological variations found in the previous study (Jiménez-Crespo, 2013a); the third group was the control group and was instructed to write from scratch directly in Spanish for the navigation options of social networking sites. The outputs were triangulated with the corpus in the previous study (ibid). The result of the experiment revealed that the translation method can influence the translated products: 'On average, selecting a translation resulted in longer and more explicit renderings while translations [produced by the first group – translations from scratch] were on average shorter and closer to the conventional items found in the corpus' (Jiménez-Crespo, 2017:155). One

limit of this study is that it only examined the translation of ‘lexical units’, investigation on the translation of ‘full sentences’ would be highly instrumental to further verify the results (Jiménez-Crespo, 2015:280). Also, because the participants in the study had reported an average of 5.62 years of experience with social networking sites, their repeated exposure to the similar text genre and its different examples of the tested lexical units may have internalised the most natural or conventionalised expressions produced in the study, which may have biased the result of the experiment.

Existing studies related to crowdsourcing translation quality focused on a limited number of text types (e.g. tourist brochure and *Facebook* interface) and translation unit (string-based). This suggests a shortage of systematic evaluations regarding the quality of ‘bigger tasks’ using crowdsourcing for different text types. My research bridges this gap by examining crowdsourcing outputs from different domains, how the crowd performs the translation in different text types, as well as how their workflows and processes can support better crowd performance.

2.5.4 Summary

This section reviewed the concept of ‘quality’ and translation quality evaluation from both academic and professional perspectives. Equivalence theory and the functionalist theories are the foundation of quality assessment, which changed as a practice from a subjective judgement to evidence-based decisions. The initial focus of quality assessment was to ensure the full transfer of source content while also respecting the conventions of target language and culture. Apart from textual analysis, it later incorporated pragmatic factors, which enriched quality evaluation to a broader level that looks at the domain, text type, and purpose of translation.

The development of professional approaches for quality evaluation demonstrates that error-based assessment is a trend which stresses ‘objective’ evaluation and assists the evaluation in a transparent and consistent way. TQA approaches developed in the 90s, e.g. Sical, SAE J2450, and BlackJack suit the traditional TEP translation/localisation era, when the content was static, and mostly business-to-customer aimed

(García, 2014). The newly developed approaches, e.g. MQM and DQF, are more flexible and take more factors into consideration when choosing the appropriate assessment model, which is more suitable in today's dynamic translation/localisation environment.

The common feature of professional TQA approaches is an error typology. As Strauss and Corbin (1998:66) indicated, the fundamental idea of classification is conceptualising and categorising phenomena according to similarities or differences. Classifications of error categories brings clarity in describing and explaining phenomena, particularly in locating errors and necessary changes in a translated text that needs to be corrected (Hansen, 2009). The assessment of translated text according to an error classification facilitates description, explanation, communication and mutual understanding, which is useful for professionals, translation trainees, and academics.

Therefore, the error-based TQA approach is chosen in this research. Through the error annotation of translated texts in different stages of collaborative translation projects (i.e. translation, self-revision, peer-revision, and the finalised output), it can not only provide a product-oriented evaluation, but also a process-based evaluation that captures how translation quality evolves under the influence of the aggregated contributions of crowdsourcing participants. Error categories and error counts offer a straightforward method to reflect on how the participants' behaviour influences the translation quality. Hence, two sources of evaluation can be achieved through error analysis:

1. Overall quality evaluation as a result of crowd contribution.
2. Effectiveness of quality control methods based on an individual's competence in translation and revision.

In my research, I investigate both aspects as they are instantiated in the crowd's contribution to real projects in different domains managed using the *Yeeyan* collaborative platform.

Chapter 3 Methodology

The intention of my PhD project is to conduct empirical research into the human and organisational features of one of the most representative Chinese crowdsourcing communities: *Yeeyan*. The research methods have been approved by the university's Research Ethics Committee (ethical approval references: PVAR 14-098 and LTSLCS-038). This chapter reviews the main methodological approaches of the study. It first delineates the purpose of each method in collecting data and how they are used, including the sampling of participants and design of the experiments. It then presents the methods used in data analysis, including the theoretical framework and the tools which have been used to facilitate the analysis.

Questionnaire survey and *observation* are the two major methods for data collection. The questionnaire survey was designed to explore the *Yeeyan* crowd's motivation for joining and the features of its participation in crowdsourcing translation. On the other hand, given that questionnaire surveys largely depend on participants using a self-assessment scale which will inevitably result in subjective data, an empirical study is also needed to examine the validity and reliability of what is 'claimed' (or 'self-reported') by the 'crowd'. Observation is a methodology that allows the collection of empirical data to explore the actual practices of collaborative projects in the chosen crowdsourced environment. It emphasises the workflow and organisation of a project as well as, very importantly, the quality of the crowdsourcing output. A *follow-up interview* was also conducted for one of the observed projects as an additional method to help interpret the observed data better.

3.1 Questionnaire Survey

3.1.1 Sampling & data collection

Convenience sampling was used to contact and select the participants. This method involved selecting participants from a population that was accessible to the researcher and potentially willing to take part in the research.

Participants for the questionnaire survey were chosen at random from the registered translators in the *Yeeyan* community. The link to the questionnaire survey was sent via the 'private-message' function on the *Yeeyan* website. The questionnaire survey was hosted on *Wenjuan.com*, which is a Chinese online survey provider. For the convenience of respondents, a local Chinese online survey website was chosen to avoid network access problems. The survey questions were presented in English with their corresponding Chinese translation for specialised terminology only. In terms of the *open questions*, respondents were given the choice to answer them either in English or Chinese.

3.1.2 Experiment design

The survey had 44 questions in total, designed to investigate what motivated and de-motivated *Yeeyan* participants, the extent of the individuals' engagement in the community, as well as their preferences for participation, methods of quality assurance, and more. Questions were designed and then grouped thematically according to the objectives of the PhD project.

There were three parts to the survey. The first part consisted of demographic questions that investigated the profile of the participants in crowdsourcing translation. The questions regarded the age group, current occupation and education background of the respondents. The second part helped investigate one of the main topics of this research – motivations, why people are willing to undertake unpaid or poorly paid translation which requires plenty of time and human effort. The third part explored people's experience and perception of their participation in a translation project in *Yeeyan*. The following sections explain the question design for the second and third part of the questionnaire (for the full questionnaire survey, please see Appendix 1).

3.1.2.1 Motivations

To gain a full-scale picture of why people engage in crowdsourcing translation, individuals were asked to answer a series of questions about what motivated and de-motivated their participation, as well as the influence of motivations on individuals' performances. The questions were designed

based on Cognitive Evaluation Theory, Self Determination Theory, and General Interest Theory.

Given the fact that previous motivation studies have shown that people are encouraged to volunteer or work by multiple factors, I chose to ask respondents to select the top 5 reasons for participating instead of selecting one primary reason. Motivation factors were presented randomly. Moreover, because other empirical studies on volunteer motivation encountered the problem of *self-awareness*, which indicates that individuals may not always fully understand their own motivations (Shye, 2009), from this point of view, neither multiple nor single choice questions were considered appropriate for this research. Respondents were therefore asked to rank their motivations in order of importance, starting with the most important one to the least important one. Ranking the selected reasons addressed the problem of self-awareness by allowing respondents to indicate the relevance of motivating factors for their individual participation. In addition, they could write down any other reason of participating if it was not among the response options.

Social desirability is another concern in survey research in general, and in motivation research, in particular. The problem is that respondents may tend to report motivations that would make them look better (ibid). The strategy of asking respondents to assess the importance of each motivation in a list prepared in advance may help alleviate the problem somewhat. However, to cope with this challenge, a question which asked the respondents to select potential motivations which he/she thought were behind other people's decision to participate in the project was also included. Moreover, participants had the possibility of suggesting additional possible motivating factors that were not included in the questionnaire.

The survey also introduced an exploratory question that investigated the de-motivation factors in order to gain a fuller understanding of the motivation mechanism in such a crowdsourced environment. The result of motivations and de-motivations can help maintain the human capital of the crowdsourcing community, and offer insights in designing an effective incentive system to promote better participation.

The relationship between motivations and individual performance is another area that the survey aimed to investigate. By asking how many hours per week one usually spent on translation-related activity in *Yeeyan*, I was able to link the work load of the person to his/her motivations. Again, the result can facilitate the design of an incentive system in the crowdsourced environment and to advise an efficient allocation of tasks.

3.1.2.2 Participation

Since the result of the questionnaire survey was expected to provide a framework of how crowdsourcing works and to provide guidance for observation, the survey included a series of questions to explore the respondents' experience with participating in *Yeeyan* translation projects.

It started from enquiring about the roles people had fulfilled to support the actual translation practice. Moreover, people's perception of crowdsourcing is a crucial area in assessing whether the crowd can follow professional ethics codes for translation. To deduce the volunteers' sense of responsibility, the questionnaire investigated the way they translate (e.g. the principles they follow), and the way they react in unexpected situation that may stop or delay their participation.

Other questions regarding the quality control methods that the respondents had used were also included. The result of the survey was used as reference in the observation stage of this research and will be compared with the output of the observation and follow-up interview.

3.2 Observation

Observation has been widely used to collect data about people, processes, and cultures in qualitative research in a variety of disciplines. Marshall and Rossman (2016:260) defined observation as 'the systematic noting and recording of events, behaviours, interactions and artifacts (objects) in the social setting.' They indicated that the term captures a variety of activities, including both being present in the setting (enacted informally) – getting to know people and learning the routines, and using strict time sampling to record actions and interactions and a checklist to tick off pre-established

actions (enacted formally). The former one refers to naturalistic observation, i.e. unstructured observation which is usually used to study the spontaneous behaviour of participants in natural surroundings without pre-determined variables. The latter refers to structured observation, which is usually conducted in a laboratory environment in which the researcher decides where, when and how the observation will take place, and uses a standardised procedure with specific variables.

Williams and Chesterman (2014:62) categorised observation as one of the subtypes of empirical research in Translation Studies: 'Naturalistic (or observational) studies are those that investigate a phenomenon or a process as it takes place in real life in its natural settings.' They pointed out that useful observational studies can be done on the working procedures of translators and interpreters. Since one of the objectives of the PhD project is to discover and identify the patterns of behaviour, interactions and relationships that occur in the crowdsourced translation projects in the *Yeeyan* community, it is better to collect data in its most natural and native settings. Hence, naturalistic (unstructured) observation was chosen as a research method.

One concern linked to using observation is that the observer may influence the behaviour of the persons who are being watched (ibid). Because the activities within crowdsourcing translation projects are all performed in the virtual environment, observation is not conducted face-to-face or in an on-site manner. The observer simply records what is going on without any intervention. The influence of the observer has thus been kept to its minimum in this research.

My observation focused on two aspects: 1. the individual's role and performance in the group; 2. the working procedures in collaborative projects. The first aspect investigated the individual's contribution in the community, including their translation activities and non-translation activities (e.g. the individual's engagement in the group interaction). The second aspect explored the workflow of collaborative translation projects in the crowdsourced environment, as well as the evolution of the translations.

3.2.1 Sampling & experiment design

To gain a thorough picture of what is going on in the crowdsourcing community, a comparative study was needed. The sampling strategy for observation was based on the comparative model of empirical research. There are two types of collaborative projects in *Yeeyan*, and observation collected data for both kinds. One is the commonly-accepted *General Collaborative Project* that people collaborate in for knowledge-sharing. There is no entrance requirement in terms of participation, and there is not much management from the organisation side, either. The other type is called the *Gutenberg Project* which volunteers join to translate books which already exist in the public domain. The purpose of this project is to publish the translations, which requires a preselection test and strict management.

One of the purposes of this project is to find out which text types are feasible for translation through crowdsourcing, and therefore a sampling strategy of observation was used with both types of crowdsourcing projects. According to the data provided by the technician group supporting *Yeeyan*, the percentage of translations produced through collaborative projects in different domains are: business – 28%, technology – 10%, culture – 15%, health – 19%, society – 13%, and media – 17% (these are the translations of the domain labels present in *Yeeyan*). Based on these results, the two most popular domains were selected for the examination of *General Collaborative Projects* – business texts and health texts. Having said that, because this is a PhD project with time and funding constraints, two more factors needed to be considered when sampling: 1. the duration of the crowdsourced projects should fit the schedule of this research; 2. the word count of the source texts for translation should be feasible for human annotation. '*Hospital Sketch*[14]' in the *Gutenberg Project* was chosen as it fits the above two considerations and also because it is a literary text, a genre that none of the existing studies in crowdsourcing translation have investigated.

With all the above factors considered, five projects were selected for

[14] 'Hospital Sketch' is a book written by Louisa May Alcott, originally published in 1863. It tells a story based on the author's experience as a volunteer nurse during the American Civil War.

observation (four *General Collaborative Projec*ts and one from the *Gutenberg Project*). Table 2 presents the characteristics of the projects based on the parameters of crowdsourcing taxonomy developed by Geiger et al. (2011).

Table 2. Characteristics of the crowdsourcing process of the observed *Yeeyan* projects

Characteristics of crowdsourcing process in *Yeeyan* Geiger et al. (2011)		
Dimensions	**General Collaborative Projects**	***Gutenberg Project***
Preselection of contributors	No	Yes – qualification based (sample test)
Accessibility of Peer contributions	Yes (assess, view and modify)	Yes (assess, view and modify)
Aggregation of contributions	Integrative	Integrative
Remuneration of contributions	No	Yes – success-based (royalties)
Settings of the observed projects		
Category	**General Collaborative Projects**	***Gutenberg Project***
Purpose of translation	Knowledge dissemination	Publication
Project management	Less governed	Strictly governed
Collaboration platform	Yes	No
Text type	Specialised text Informative text type	Literary text Expressive text type

The first dimension, ‘preselection of contributors’, is concerned with restrictions regarding the group of potential contributors (ibid). The *General Collaborative Projects* in *Yeeyan* do not limit the number of contributors, and any registered member can volunteer for a task when a project is initiated with an open call. By contrast, for the *Gutenberg Project*, because the purpose of translation is for publication, a higher level of quality is expected. Contributors are filtered based on their qualification. The organiser (a recruited text editor in *Yeeyan* or a specifically selected project manager) usually picks a passage from the source text and asks the potential contributors to submit their translation within a deadline. The organiser then

chooses suitable translators for the project.

The second dimension, ‘accessibility of peer contributions’, indicates to what extent contributors can access each other’s contribution. There are four characteristics of this dimension which reflect the degree of access that a crowdsourcing process enables. They are, in order of increasing accessibility: none, view, assess, and modify (ibid). In *Yeeyan*, both types of collaborative projects are accessible for team members, meaning that participants in a project are able to view, comment on (i.e. assess) and modify each other’s translation. The difference between the two types of projects is that ‘accessibility’ is only available for the team members within the *Gutenberg Project* but is publicly available for all *Yeeyan* members in a *General Collaborative Project.*

The third dimension, ‘aggregation of contributions’, describes how the crowd contributions within a crowdsourcing process are used by the crowdsourcing organisation to achieve the desired outcome (ibid). The crowdsourcing process in *Yeeyan* is designed around ‘integrative contributions’, meaning that all contributions are encouraged and used for the final outcome. Another type of this dimension is ‘selective contributions’ which indicates that individual contributions are compared to each other and the ‘best’ one(s) is selected, e.g. crowdsourcing translation in *Facebook* (Mesipuu, 2011).

The fourth dimension, ‘remuneration for contributions’, determines how contributors are paid or otherwise compensated their work (ibid). Generally, it includes fixed remuneration (contributors receive a fixed payment regardless of the value they add to the final product) and success-based remuneration (contributions will be paid depending on their individual value to the crowdsourcing goal). In *Yeeyan*, contributions in *General Collaborative Projects* are completely voluntary, meaning that there is no payment involved. For the *Gutenberg Project*, contributors receive a portion of royalties depending on the sale volume of the book that they translated.

As shown in Table 2, according to the taxonomy of crowdsourcing process developed by Geiger et al. (2011), ‘preselection of contributors’ and ‘remuneration for contributions’ are the two dimensions that differentiate

General Collaborative Projects from the *Gutenberg Project* in *Yeeyan*. In an empirical study, these two dimensions can be used as variables to compare these two different processes. However, as there is no direct measurement that compares people's performance with the remuneration that they receive, this research uses the first dimension as one variable – to compare the competence of general participants and the preselected participants in *Yeeyan*. Moreover, the *high* 'accessibility of peer contributions' and *integrative* 'aggregation of contribution' are two evident characteristics of crowdsourcing translation in *Yeeyan*, which differentiate it from other crowdsourcing practices (for further details, see section 2.2.1, Literature Review). The data from the observation stage will be used to examine the effectiveness of these two *dimensions* in the *Yeeyan* crowdsourcing process. A comparison will also be made in terms of the *different settings* of the two types of projects in *Yeeyan* (Table 2).

Table 3 presents some features of the selected projects for observation (for the full specifications, see Appendix 2). First of all, the source texts of the *General Collaborative Projects* are press reports in specialised domains, while the source text of the *Gutenberg Project* is a classic literary text. *Number of translators* refers to the participants who initially signed up for the translation task in each project. *Average word count for each translator* is determined by the project manager who divided the source text according to the number of translators.

Table 3. Features of the observed projects

Project Type	Project Name	Text Type (domain)	Text length (in words)	Number of translators	Average word count for each translator (*approximate*)
General Collaborative Projects	*Health Project I* (*HP I*)	Health	10,607	11	960
	Health Project II (*HP II*)	Health	5,389	7	760
	Business Project I (*BP I*)	Business	6,080	8	760
	Business Project II (*BP II*)	Business	962	3	360
Gutenberg Project	*Gutenberg Project* (*GP*) (book title: Hospital Sketch)	Literary	28,886	2	14,400

3.2.2 Data collection

There are two sources of data for observation: one from documentary material (*General Collaborative Projects*) and the other from a large on-going translation project (*Gutenberg Project*). Documentary material refers to the translation projects which take place on the online platform. The collaboration platform is available for open access and stores the entire translation process, including the participants' edit history and their interaction record with a clear timeline. Below are two anonymised representative screenshots of the collaboration platform and the chat board interface.

Figure 10. Screenshot of the collaboration platform for the *General Collaborative Projects*

The four *General Collaborative Projects* were performed via *Sync.In* – a web-based collaborative word processor allowing real-time synchronisation of multiple edits. As can be seen in Figure 10, it uses different colours to label different participants' work (for privacy reasons, the information relating to the participants' identity is blurred). Edits are saved automatically and can be viewed through the 'Time Slider' option on the top right of the interface. The interaction board is located on the right bottom section of the platform interface. It is a built-in text-based live chat board for project participants. The volunteers' presence and activity in the chat board are marked by changes in colour, and the archives of a conversation can be accessed using the scroll bar. One can always revisit and check the communication history by scrolling up or down.

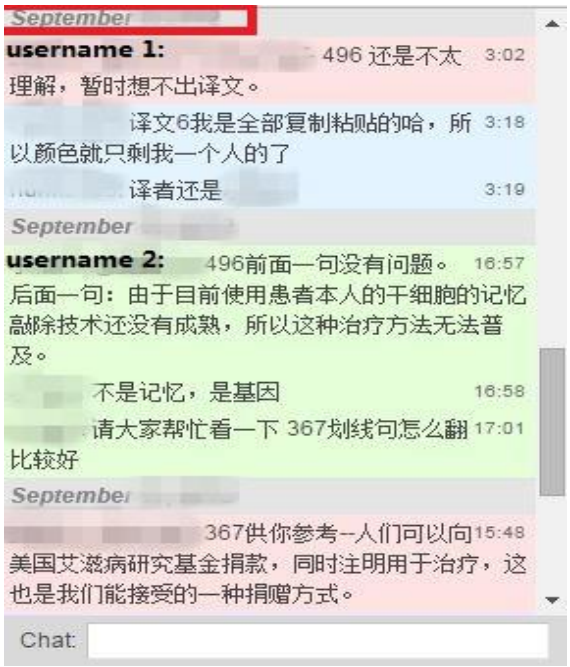

Figure 11. Screenshot of the built-in chat board in the collaboration platform for the *General Collaborative Projects*

Participants choose a colour and username to represent their identity in the project upon logging in (highlighted in the red rectangle in the top right part of Figure 10. The online presence of participants is updated in real-time and shown in the red circled area in Figure 10. It helps recognise who joins the project and who is contributing at any time. One can type messages in the chat box (highlighted in the red rectangle in the bottom right part of Figure 10), which are then displayed sequentially on the chat board. These messages convey two types of information: one is the content of the message, the other is the presence of the participant. Messages are shown in the form of chat lines and are labelled with the sender's username and sending-time (Figure 11).

The *Gutenberg Project* was not conducted through the publicly-accessible collaboration platform because of its particular characteristics. Participants use *Microsoft Word* to translate and the 'group chat' function in *Tencent QQ* (a Chinese instant-messaging software) to communicate with each other. Translations and other related files are uploaded to the repository of the chat group for sharing.

Follow-up interviews were conducted after the finalisation of the *Gutenberg Project* (for the interview questions, please see Appendix 3). It was used as a supplement to provide explanations for the data collected from observation. I conducted three individual interviews with the participants (project manager and two translators) in the *Gutenberg Project* (for details, please see section 5.4). Each interview lasted around 30 minutes, mainly exploring the person's motivations and his/her explanation of their behaviour in the group, the problems they had encountered, as well as what they had achieved through participation. More importantly, the interview investigated the tangible benefits which participants had gained throughout the translation project. The interview results will be mentioned when discussing the observation data. The reason why there was no follow-up interview for the four *General Collaborative Projects* is because the data are archived material. After the update of the *Yeeyan* website, it was not possible to trace the participants' identity and contact them.

During the observation stage, I followed the interactions between participants which were automatically saved in the communication tools that these projects used. For *General Collaborative Projects*, I used the '*Time Slider*' function of the collaboration platform to retrieve every version of their outcome from translating and revising, until publishing. For the *Gutenberg Project*, translation texts and other related files were collected from the group's repository when it was ready for download. The aim of using observation to collect data is to track interactions and changes both in terms of collaboration and in terms of the translated product during production, and then to produce a quality evaluation of the final product, as well as an evaluation of the participants' performance. In short, with the explanation of how observation can be done in the research, two types of data were collected:

- The communication record of each project
- Translation texts and participants' edit history of each project

3.2.3 Data analysis 1: Communication data

3.2.3.1 Step 1: Dialogue act analysis

The first step in data analysis was to annotate the communication record of each project to find out what people had been talking about while participating. When annotating the communication record, an interaction typology with specific dialogue acts was developed. In light of the studies in interaction in collaborative systems (Soller, 2001; Babych et al., 2012), dialogue acts are defined and grouped thematically as:

- General interaction – acts which bear the task assignment of the project, promote the flow of the dialogue, and fulfil the need of social networking between participants;
- Information – acts which bear on the provision and discussion of information related to the project and the substance of the translation;
- Maintenance – acts which signal the reception of information, appreciate and encourage member's participation;
- Status – acts which bear on the progress of the work schedule.

Table 4 explains how interactions are categorised and defined in this research.

Table 4. Interaction typology in the study

Interaction Category	**Dialogue Acts**	**Description**
General Interaction	Commission (CM)	To inform a task; to claim a task.
	Comment (CT)	To comment on the source text (ST) or topic related to the ST; to comment the progress of the project (e.g. speed, work flow, etc.); to comment on the last request (without giving direct answer).
	Social (SO)	To socially network with peers; to discuss things that are not related to the project.
	General information	To provide general information

Information	(GIF)	related to the project (e.g. translation requirements, project schedule, platform usage, etc.).
	Information request (IR)	To ask technical questions (e.g. use of the platform) and project-related questions (e.g. work flow) – typically a 'WH' question.
	Information clarify (IC)	To provide answer for the last information request.
	Term request (TR)	To ask questions that relate to the translation of a specific term.
	Term clarify (TC)	To explain or directly provide a translation for the requested term; to argue or justify one's reason for the translation choice of a given term.
	Text request (XR)	To ask questions that relate to the translation of a sentence, a paragraph or a part of the source text.
	Text clarify (XC)	To explain or directly provide a translation for the requested text; to argue or justify one's reason for the translation choice of a given text.
	Revision request (RR)	To ask someone to revise one's translation.
	Revision comment (RC)	To provide revision comments or suggestions for one's translation on a general level.
Maintenance	Motivation (MT)	To inspire group members; to encourage others' involvement.
	Acknowledgement (AK)	To signal to peers that one has received the information; to show appreciation for other's comments on or contributions to a task; to compliment one's work.
Status	Status check (SC)	To ask about progress on a specified task.
	Status report (SR)	To respond to a check on progress on a specified task; to proactively report one's progress.
Miscellaneous	Miscellaneous	Any dialogue act that does not fit in

	dialogue act (MIDA)	the above categories.

Here, it is necessary to point out that, in the categories of dialogue acts, ‘term’ in term request (TR) and term clarify (TC) includes both specialised terminology and general glossary. Term request (TR) represents any request for information on the lexical level, and text request (XR) represents any request for information on the sentence or above sentence level (e.g. one participant may ask another to check a sentence where he/she has problem, or may have trouble in translating a whole paragraph and therefore requires help).

When annotating the collected communication record of each project, three procedures were used:

1. Extract communication information in the following fields: text – content of each dialogue; source – sender of the message (identity and role); target – receiver of the message (identity and role); timestamp – the exact time of when a message was sent.
2. Categorise and label the dialogue acts according to the interaction typology (stated in Table 4).
3. Compute the categories and frequency of dialogue acts.

Participants’ dialogue acts were counted based on the dialogue thread. The beginning of a new dialogue thread was identified by the start of a new context or topic because direct addressing was not always used (or required) to specify the target of a message. A message without explicit direct addressing was assumed to target everyone in the group.

Based on the above interaction typology, the first step of analysing the communication data was to annotate and calculate the categories and frequency of dialogue acts. The aim was to establish indicators and criteria for evaluating interaction in crowdsourcing translation projects, and to provide context for the next step of data analysis which will help examine the power relationship between participants connecting to the dialogue acts.

3.2.3.2 Step 2: Social network analysis

The second step in analysing communication data was the social network analysis (SNA). On the one hand, SNA investigates the relationships and information flows between the participants in the observed projects; on the other hand, the commonly used strategy in SNA – visualisation – can be used to interpret the collected data explicitly.

As defined by Marin and Wellman (2011:11), 'A social network is a set of socially relevant nodes connected by one or more relations. Nodes or network members are the units that are connected by the relations whose patterns we study. These units are most commonly persons or organisations, but in principle any units that can be connected to other units can be studied as nodes'. The task of the researcher is to describe and explain the patterns exhibited in these connections (i.e. the interactions between the participants).

Social network researchers measure network activity for a node by using 'degrees' – the number of direct connections a node has. 'Centrality degree' and 'density' are the two metrics when analysing the networks in this research. *Centrality degree* refers to the number of connections that a node has (Kosorukoff, 2011). It is used to identify the 'key' actor in each network (i.e. the most active participant in group interaction). For interaction data, as one string of the dialogue contains a source and target of the message, a connection between the nodes is made. There are two separate measures of centrality: in-degree and out-degree. Here, in-degree is determined by the number of connections directed to the node, i.e. the number of messages sent to one participant; out-degree refers to the number of connections that the node directs to others, i.e. a count of messages that are sent by one participant. In this study, *weighted out-degree centrality* was used to measure the importance of actors in the network in each project. Dialogue acts regarding the provision of project-related information (general information-GIF, information clarify – IC) and knowledge exchange of translation-related messages (term request – TR, term clarify – TC, text request – XR, text clarify – XC, and revision comment – RC) were double-weighted compared with other dialogue acts, as they directly influence the progress and quality of

the translation. Hence, weighted out-degree centrality was determined by both the number and type of dialogue acts the participant performed in the chat board.

Density, refers to the proportion of connections in a network relative to the total number possible (ibid). It was used to compare the networks of observed projects to find out how closely the participants are connected in each project through interaction. Through social network analysis, the participants' performance in both individual level and collective level can be compared.

Visualisation is the common tool used to present the features of a social network. The advantage of social network visualisation is that it reduces the complexity and enables the researchers to pinpoint easily key participants and clusters within the network, by using a variety of metrics from social network analysis as the foundations for visualisation.

There are a variety of social network visualisation packages on the market – open source and commercial. *Gephi* was used in this research. As an open source software, it has a large pool of active and highly responsive technicians and user community members to help solve problems in utilisation. More importantly, as opposed to other tools, such as *Pajek* (open source) and *Ucinet* (commercial), it is more user-friendly for analysis and visualisation. Many social network visualisation tools only process matrix data as source input, which requires the user's knowledge in statistics and takes more time and effort in converting the data into matrix format. *Gephi* can import a simple CSV file as source data as long as it includes the source and target (actors) of a connection (relationship) in its readable format. Other attributes could be included as well (e.g. label of the interaction category and timestamp).

A concern with social network visualisation is that it can overshadow the analysis which merely converts data into a graph (Bruns, 2012). Hence, it is important to develop a clear understanding of the implications resulting from network visualisation. To avoid this problem, I also considered the result from the dialogue act analysis when annotating the social network. The analysis

explores what kinds of dialogue acts make a 'key' participant in a network, their influence on the project as a whole, as well as the difference in terms of the role of the 'key' participant.

3.2.3.3 Step 3: Comparison with Students' Team Project

The third step in analysing interaction data was to compare the result from the *Yeeyan* crowdsourcing projects with the *Students' Team Project* in the Computer-Assisted Translation module of the *University of Leeds MA in Applied Translation Studies* Programme. The *Students' Team Projects* mirror the industry localisation ones to give students the opportunity to gain a real sense of the challenges which translation professionals experience in their daily work, including communicating with colleagues and clients, planning and organising the translation or localisation project, collaborating between team members, and fine-tuning the relationship between the translator, the agency and the client. I gained access to the communication data in the students' team project with approval from the module leader and the university's Research Ethics Committee.

The reason for such a comparison was that both the *Students' Team Projects* and the collaborative projects in the *Yeeyan* crowdsourced environment resemble to a large extent the real-life translation practice in the professional industry as observed during my visit to *Sandberg Translation Partners Limited* (STP), one of the largest Language Services Providers in the UK. In the first case, students work with different language combinations and translate individually within a project, whereas in the second setting, i.e. the observed projects in this study, volunteers work with English-Chinese collaboratively, and each of them performs a part of the translation work for one source text. The final product is a collective output within a project.

The *Students' Team Projects* are designed with input from major players in the Language Services world such as *Google*, *Welocalise* and *Sandberg Translation Partners Limited* to mirror multilingual translation projects in industry, while the observed crowdsourced projects in this research are also similar to single-target-language industry translation projects that only have one large source text. Reviewing the literature in social communication

studies, it is surprising that there appears to be a relative dearth of communication analyses in the translation domain. This is probably because, on the one hand, it is difficult to collect data from the industry; on the other hand, the role of communication has been somewhat underestimated in both translation practice and translator training, although it is listed among the important new skills which translation graduates need to have in order to enter the Language Services Industry (ELIA et al., 2017). Consequently, this comparative study focusing on the communication aspect of collaborative projects in both crowdsourced and industry-like environments is expected to fill the gap in communication studies in the translation domain.

A major difference when comparing *Students' Team Projects* with crowdsourced projects lies in the traits of actors (translation competence, personal preference in communication, etc.) and duration of the network (schedule of the projects). The former set of features may influence the effectiveness and the frequency of interactions between actors, while the latter may influence the total number of communication acts in the projects. To address this bias, SNA appears to be the most suitable method of research because it focuses on the relationships/connections between actors rather than the properties of actors. After performing a dialogue act analysis together with a social network analysis, I then compared the communication patterns in the two environments by assessing the structure of the network, the category and frequency of core dialogue acts, and the characteristics of the 'key' actor. The result of the comparison was used to establish a comprehensive set of communication features in collaborative translation projects, and to offer insights on designing effective collaborative projects in both classroom and crowdsourced environments.

3.2.4 Data analysis 2: Translation texts and participants' performance

An archive of translated texts and participants' edit history was collected through observation. With the purpose of assessing the quality of crowdsourcing translation, data analysis focused on two aspects: the workflow of the two types of collaborative projects in *Yeeyan* and the features

of the translation outputs.

3.2.4.1 Step 1: Source text analysis

Prior to the error annotation of the translated texts, a source text analysis was conducted to determine the difficulty level of each project and subsequently help evaluate the competence of participants more fairly. The difficulty level of the source texts and the purpose of the translation were two important parameters later considered in order to give a holistic evaluation of translation quality.

A readability score and the amount of specialised terminology in the source texts were the two factors in estimating the difficulty level of each project. The *Flesch Kincaid Reading Ease*[15] is a commonly-used formula for evaluating the readability of texts written in English, calculated using the number of syllables in 100 words and the average sentence length (Humphreys and Humphreys, 2013). Based on a scale of 0-100, a high score means the text is easier to read, while a low score suggests the text is complicated to understand. The *Reading Ease* score is not a fool-proof mechanism in Translation Studies, because it cannot anticipate how familiar the translator is with the source text or domain terminology, regardless of how short the words themselves are, but it is nevertheless an indicator of potential difficulty of the source text given that the longer a sentence is, the more difficult it can be to resolve anaphoras, as well as theme-rheme identification.

For the five observed projects, the source texts of two projects were assessed as 'fairly difficult' (*Business Project I* and *Gutenberg Project*). *Health Project II* was labelled as 'difficult', while *Health Project I* and *Business Project II* were assessed as 'normal'.

However, given the shortcomings of the *Reading Ease* test, another readability test was also used – *McAlpine EFLAW*[16], which reflects the

[15] The formula for the Flesch Reading Ease Score was: 206.835 – (1.015 x ASL) – (84.6 x ASW), where ASL = number of words divided by number of sentences and ASW = number of syllables divided by number of words (Dalecki et al., 2009).

[16] The formula of McAlpine EFLAW Readability Score is (W+M)/S, where W stands for the number of words, M stands for the number of miniwords, and S stands for the number of

difficulty of any given text for non-native English speakers. The formula is based on the proportion of long sentences and miniwords[17]. It is believed that the two elements are confusing for ESL (English as a second language) learners as they may have ambiguous meanings and are often a sign of wordiness and idioms (McAlpine, 2005). In the *McAlpine EFLAW* test, the source text of the *Gutenberg Project* was assessed as 'very confusing'; *HP II* was in the second category as 'a little difficult to understand'; whereas *HP I* and *BP I* were assessed as 'quite easy to understand'; and B II was 'very easy to understand'.

Table 5. Source text analysis (readability scores, number of specialised terms, and difficulty level)

Projects	Flesch Kincaid Reading Ease score	Difficulty 1	McAlphine EFLAW score	Difficulty 2	Number of specialised terms	**Final Difficulty Level**
HP I	66.4	normal	23.9	quite easy to understand	66	**difficult**
HP II	42	difficult	27.9	a little difficult to understand	52	**difficult**
BP I	56.2	fairly difficult	24.3	quite easy to understand	63	**difficult**
BP II	64.5	normal	19.5	very easy to understand	12	**normal**
GP	57.6	fairly difficult	45.9	very confusing	3	**very difficult**

Table 5 presents the readability score under each formula. As the two tests use different criteria for the evaluation of difficulty, it is understandable that we had varied results. However, they both suggest that *HP II* and *GP* are challenging texts for both native and non-native English speakers. According to these measures, the source text of *BP II* is a relatively easy piece of reading, which corresponds to the initial purpose of *BP II* – to serve as translation practice for *Yeeyan* beginners. Moreover, *HP I* and *BP I* are

sentences in a text. The resulting score can fall into one of four categories: very easy to understand (1-20), quite easy to understand (21-25), a little difficult to understand (26-29), and very confusing (30+) (Kraut, 2017).

[17] Miniwords refer to short, common words of one, two or three letters.

somewhere between normal and easy.

Having evaluated the source texts according to two popular readability scores, one still cannot ignore the fact that translation is an activity that not only depends on the translator's language competence, but also on knowledge of the source text domain (thematic competence), and many other competences (EMT expert group, 2009). Readability scores based on word and sentence lengths cannot fully represent the difficulty level of translation in each project yet.

To give a more realistic picture of the level of difficulty of the chosen source texts, the density of specialist terminology was also measured. The number of specialised terms reported in Table 5 is based on the glossary list extracted by the participants themselves during the translation task combined with a supplementary one extracted by the researcher using the *Keywords* function in *Sketch Engine*. During the error annotation of the translated texts (step 3 in the quality assessment), I found that the glossary list did not encompass all the errors made in the category of specialised terminology. Hence, a supplementary one is needed to provide a comprehensive evaluation of the source text.

Therefore, considering the topic of the source text, the amount of specialised terminology and complex sentences, and the text length, together with the readability scores, I drew the conclusion that, for non-professional translators, the translation of *GP* is a 'very difficult' task. As a classic literary text, it has few specialised terms. However, as reflected in the readability score of *McAlphine EFLAW*, it has many complex long sentences and miniwords which is seen as 'very confusing' for non-native English speakers. Moreover, the translation of literary texts does not only require accuracy and fluency, but also needs to maintain the source text's aesthetic features. This text was therefore considered to be the most difficult one among all the observed projects. Three of the *General Collaborative Projects* (*HP I, HP II,* and *BP I*) are 'difficult' tasks due to the relatively higher number of specialised terms.

BP II is estimated as a 'normal' task, mainly because of its short text length,

the easy level of readability score, and the amount of specialised terminology.

3.2.4.2 Step 2: Workflow mapping and productivity analysis

To begin with, the participants' edit history indicated the type of linguistic activity individuals performed in the observed collaborative projects. When annotating the linguistic activity, I first distinguished the types of edits individuals made – from translation to revision (self-revision and peer-revision), and then also logged the time the individual spent on each activity.

The annotation of the linguistic activity was visualised in order to map the workflow of each project. In the visualisation graph, major workflow stages were represented along the *Y* axis, and the individuals' activity in each stage was represented on the *X* axis in chronological order – see Figure 20, page 152. Through workflow mapping, I was able to compare working procedures with regard to the influence of the degree of project management on the participants' behaviour – the differences between less governed and strictly governed projects. The workflow of each project is essential when assessing the translation quality in order to make suggestions for optimising the workflow and organisation of collaborative projects in the crowdsourced environment.

In terms of productivity assessment, two statistical measures were used: mean and standard deviation. First, the average productivity during the two linguistic activities – translation and revision – was calculated. I then compared the average productivity in the observed projects with the average productivity expectations of the professional industry and translation trainees' productivity collected from the *Students' Team Project* in order to examine the validity of the assumption that crowdsourcing produces 'faster' translation output.

3.2.4.3 Step 3: Error annotation and quality assessment

The quality of the translated texts is the most straightforward measure of the effectiveness of collaborative projects in the crowdsourced environment. Error-based assessment is a commonly-used approach in evaluating translation quality in the professional industry, which uses translation quality

assessment (TQA) as a more objective, data-driven method instead of the traditional subjective analysis of linguistic features. Research has also confirmed the validity and reliability of using error analysis to evaluate translation quality (Kunilovskaya, 2015; Waddington, 2001). Describing the collected translated texts according to the error categories they feature at different stages of the collaborative projects can offer a more methodologically-valid approach to offer feedback, and more importantly, to provide suggestions for improving the crowdsourced output. Hence, error annotation was chosen to assess the translation quality and the participants' performance in the observed projects.

The error typology used in this project was based on the one developed during the EU-funded *MeLLANGE* project, revised for marking students' translations by Dr Serge Sharoff, and further revised by me to cater for the specificities of *Yeeyan* – namely, I added an error category for stylistic errors (*literal translation* and *inconsistency within the text*) that occurs in the crowdsourced environment possibly due to the lack of governance and professional quality control mechanism (for other reasons for using this typology see page 75, Literature Review). Table 6 presents the error typology that was used for quality assessment in this project.

Table 6. Error typology in the study

Category	Sub-Category	Category	Sub-Category
Content errors	Omission of information (OM)	**Language errors**	Ambiguous translation (AM)
	Addition of information (AD)		Awkward syntax (AW)
	Meaning distortion (DI)		Not idiomatic, non-native use of target language (NN)
	More suitable term available (TE)		Problems with register (RE)
	Term not correct (WTE)		
Style errors	Literal translation (LT)	**Hygiene**	Incorrect punctuation (PU)
	Inconsistency within the text (TI)		Character Spelling (CS)
Miscellaneous	(MI)		

A computer-assisted translation (CAT) tool – *SDL Trados Studio 2015 Professional* – was used to facilitate the error annotation. Advantages in using the TQA function in *SDL Trados Studio* are:

1. It supports the creation of customised TQA model (i.e. the error typology in Table 6);
2. The CAT environment improves annotation productivity and helps ensure the consistency of the assessment (e.g. one does not need to remember the error typology by heart);
3. It generates a bilingual TQA report which contains scores, the used error typology, and 'track changes' of the annotated translation, which makes the follow-up statistical analysis time-effective and consistent.

Three versions of the translation outputs in each project were assessed: the translation, the self-revised version, and the peer-revised final version. The identification and annotation of translation errors was mostly done by comparing the version to annotate with the final published translation, with occasional input from my own expertise, as well. To minimise the number of

instances of researcher bias in evaluation, when annotating the errors, the final published version was considered the ‘gold standard’ because it is the result of collective effort and is well-accepted by the readers. There were only very few cases of translation problems still present in the final published version that were identified using my own knowledge and experience in translation studies.

By doing error annotation of the three versions of translation texts, I aimed to investigate:

1. The effectiveness of quality control methods in a crowdsourced environment (self-revision and peer-revision);
2. The competence of crowdsourcing participants in performing linguistic activities (translation and revision);
3. The translation quality produced by crowdsourcing.

Error analysis focused on two aspects: the ‘frequency’ and ‘category’ of errors made by participants at different stages of the observed projects. The result was used to examine the effectiveness of quality control methods in the crowdsourced environment (self-revision and peer-revision). Correction rate was one of the parameters in evaluating the quality of translated texts and effectiveness of the organisational set-up, as it reflects the shaping of translation outputs through the aggregation of contributions.

Error analysis also looked at the significant error types that the participants made, as well as whether they were corrected through revision. To evaluate the participants’ translation competence, mean and standard deviation were the two statistical measures that helped identify the common error types made by participants in the crowdsourced environment. The result will help find out the particular knowledge and skills crowdsourcing translators lack.

The T-Test is the statistical method that helped identify the significant error types. By comparing the error count under each category between the draft translation and the revised translations, the result highlighted significant differences made by revision, which further pointed out the significant error types that participants should correct.

Further analysis investigated whether there were any differences in terms of the significant error types in different text types, as well as the level of competence of general participants and preselected participants.

To evaluate the competence of crowdsourcing participants from the perspective of revision, a correlation analysis was conducted to find out if participants have the ability to correct a similar amount of errors and error types in others' translations compared with the errors that they make in their own translations.

The quality assessment of the final translation was based on the Chinese industry criteria, given that this PhD project focuses on crowdsourcing practices in China. According to the Specification for Translation Service – Part 1 (GB/T 19682-2005, GB/T19363.1-2008) announced by the State Administration for Quality Supervision and Inspection (2005, 2008), four factors should be considered when assessing the quality of the target text: 1. purpose of translation, 2. genre, style and quality of the source text, 3. difficulty of the text domain, 4. time-frame of the project. It specifies a method of assessment by calculating its overall error rate:

Overall error rate = K*CA*[(CI*DI+CII*DII+CIII*DIII+CIV*DIV)/W] *100%

- 'K' refers to the overall difficulty level of the project, and the recommended value range is 0.5-1.
- 'CA' represents the coefficient for the translation purpose. The recommended values are:
    - a. translation for formal documents, legal file or other texts for publication, CA=1;
    - b. translation for general materials and documents, CA=0.75;
    - c. translation that will be used as reference material, CA=0.5;
    - d. translation for gisting, CA=0.25.
- 'CI, CII, CIII, CIV' represent the coefficient for the category of error:
    - a. semantic error of the key content, meaning distortion of critical words (numbers), omission of sentences and paragraphs, CI=3;

    - b. general semantic error, meaning distortion of non-critical words (numbers), grammar mistake, inaccurate word choice, and ambiguous expression in the target text, CII=1,
    - c. inaccurate translation of specialised terminology, inconsistency of translation, and failure to meet target language convention, CIII=0.5;
    - d. failure to meet rules and requirements in translating measure word, symbols, and abbreviations, etc., CIV=0.25.
- 'DI, DII, DIII, DIV' refer to the error count under each defined category. Repeated errors are counted as one.
- 'W' stands for the number of words.

The overall error score shall not exceed 1.5‰ (any text shorter than 1000 words is counted as 1000 words).

To assess the final translation quality, the aforementioned evaluation approach in the Chinese professional industry was used. The difficulty estimation calculated in Step 1 of analysing translation texts was used when determining the 'K' value for the formula. The result of the overall error rate estimation in each project will indicate whether the final translation output meets the quality requirement in its particular domain.

3.3 Summary

This chapter explained the methodological considerations of the PhD project. The questionnaire survey, observation and follow-up interview were the three main methods for data collection. The *questionnaire survey* was designed to give a full-scale picture of crowdsourcing in the *Yeeyan* community. It involved questions that explore the demographics of the community, motivation, and their actual participation (i.e. how people engage in translation-related activities and how quality is controlled.). From the perspective of practitioners in *Yeeyan*, based on their own perceptions and experience, the survey outcome is expected to, first of all, de-mystify the 'motivation' side of the research questions. Secondly, the result is expected to form a general framework in analysing empirical data collected in this

domain and type of collaborative environment.

Through *observation*, empirical data on the collaborative projects was collected, including crowdsourced projects with different settings (e.g. purpose of translation, the degree of project management, preselection of participants, and text type). Two types of data were collected: the interaction record for each project, and the translated texts and participants' edit history. The former type of data was analysed using *dialogue act analysis* and *social network analysis*, which helped find out the role of interaction in collaborative projects in a crowdsourced environment. The latter type of text data was analysed using *workflow mapping* and *error annotation*, which will help assess the translation quality and the participants' competence in performing linguistic activities (translation and revision).

The *follow-up interview* in one of the observed projects (the *Gutenberg Project*) was used to supplement the text data.

Finally, a data analysis of all collected materials informed the discussion of the validity and reliability of using collaborative projects in the crowdsourced environment and offered insights into designing and managing effective crowdsourcing translation practices.

Chapter 4 Motivations and self-declared participation

This chapter reports the findings of the questionnaire survey. As mentioned in the methodology chapter, an invitation to participate in a Web-based survey was sent via the 'private message' function in *Yeeyan* to members. 150 messages were sent, and 62 valid responses were received. A response rate of 41.3% was achieved.

According to Shye (2009:187), there are three classes of causes that determine why people act in a certain way and choose to volunteer: demographic characteristics ('the personal resources and assets one must have in order to *be able to* volunteer'), motivations ('more specific and necessary [reasons] for the volunteer in order *to want* to volunteer'), and circumstances ('triggers or accidental events, that *prompted or enabled* volunteering'). In the order of Shye's classification, this chapter first discusses the demographics of the community, particularly the education background and language competence. It then elaborates on the motivations and other factors that may influence the participation in crowdsourcing translation projects. The final part reports the findings regarding the actual 'participation' in translation-related activities in *Yeeyan* compared with the respondents' own perceptions and experience. This is done in order to provide a reference for the analysis of empirical data in the following chapters.

4.1 Respondent Demographics

The vast majority of the respondents were between 20 and 30 years old. Although 95% of the respondents had college degrees or higher, a small (5%) proportion were high school students. Nearly half (42%) of the respondents were majoring or had majored in language-related studies, and 30% received translation-related training. As indicated by the respondents, the translation-related training they had included bachelor/master in translation and/or interpreting, literary and business translation courses, localisation, and project management.

Regarding the current occupation of the respondents, 33% of them were students, 23% were engaged in language-related work, while 37% were in other areas, such as civil service, management, and information research. An interesting fact was that 7% of the respondents were professional translators.

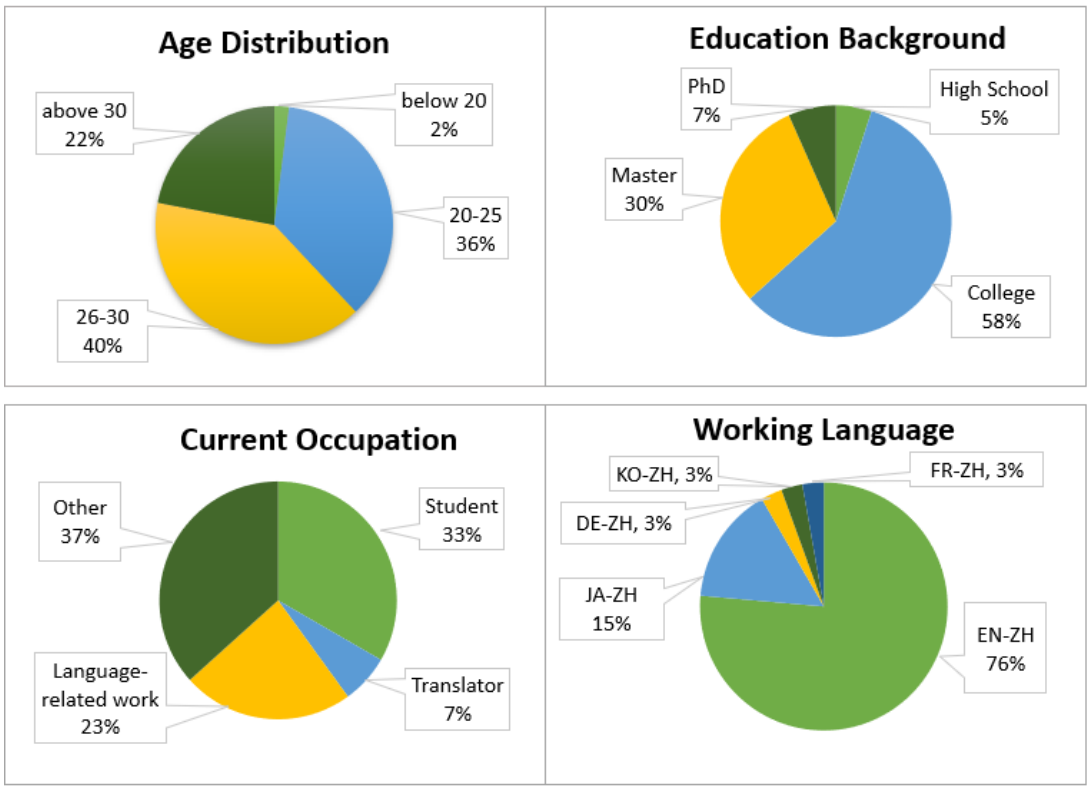

Figure 12. Respondents characteristics

Figure 12 shows the general characteristics of the survey respondents. The main reported working languages in *Yeeyan* were English-Chinese (76% of the respondents). Other language combinations were Japanese-Chinese (15%), German-Chinese (3%), Korean-Chinese (3%), and French-Chinese (3%). It indicates a diversity of language competence among *Yeeyan* members.

4.1.1 Language Competence

Yeeyan members are usually native Chinese speakers. In order to perform translation activity, they must be L2 learners at some stage. According to Harris (1977) and Whyatt's (2012, 2017) view in natural translation, L2 learners are able to translate but their performance varies depending on the

stage they are at along the developmental continuum of human translation ability. Since this research aims at investigating EN-ZH translation in the crowdsourced environment, the survey explored the English (L2) competence of the respondents by asking the level of English language certificate that they had. Results show that, for those who mainly engaged in EN-ZH translation, respondents usually had more than one English language certificate.

Commonly used certificates for English competence in China are: CET 4 and 6 (College English Test Band 4 and Band 6), TEM 4 and 8 (Test for English Majors Band 4 and Band 8), IELTS (The International English Language Testing System), and TOEFL (The Test of English as a Foreign Language). CET 4 and 6 are national standard tests to measure the English ability of college students so as to offer insights for English teaching in higher education institutions. It is a criterion-related norm-reference exam including listening, reading, writing, and translation tests. CET is widely recognised in the Chinese society and has become an employment criterion for college graduates. TEM 4 and 8 are also national standard tests but are aimed at English majors and English-related majors. IELTS and TOEFL are international tests that measure the language proficiency of people who want to study or work in environments where English is used as a language of communication.

In the survey, I divided the domestic exams (CET and TEM) into two groups: *pass* and *merit*, whereas I grouped the two international exams according to the general level of international college entrance criteria. For instance, IELTS was divided into two groups: 6-6.5, and 7+ (6-6.5 is the basic entrance requirement for Master programmes in the UK, while 7+ is the entrance requirement for Master programmes in translation and/or interpreting studies in the UK); the new TOFEL iBT (internet-based test) was also divided into two groups: 80-90, 90+ (the average entrance requirement for Master degree in the United States is 80, 90+ is the entrance requirement for top liberal arts colleges in the United States). In this research, TEM *merit*, IELTS 7+, TOFEL 90+ were regarded as the advanced level of English proficiency. CET *merit*, TEM *pass*, IELTS 6-6.5, and TOFEL 80-90 were regarded as the

intermediate level; CET *pass* was regarded as the basic level.

While these English test certificates cannot fully represent one's language competence and such a division of language level lacks sufficient data support, they are still useful given that the subjects of this research are in an open entrance environment which makes it very difficult to set a fair and acceptable language test for *Yeeyan* members. By asking what kind of language certificates they possess, I could, at least, gain a basic understanding of the English level of this online community. How it compares to the actual quality of translation produced from this community is discussed in 6.4.

The survey result shows that a quarter of the respondents had TEM certificates, indicating the graduate level of English/English-related majors. 28% of the respondents were in possession of international standard certificates (IELTS and TOFEL). As shown in Figure 13, English competence among the respondents is almost evenly distributed: 26% of the respondents can use English at a basic level, 33% are at the intermediate level, and 36% are at the advanced level. Three respondents declared that they had none of the certificates listed on the survey. Two respondents defined their competence as 'expert' and 'near native' in the comment area of the question. The survey result implies that a certain group of members in the *Yeeyan* community are able to use English proficiently or at least at a similar level as the trainee translators.

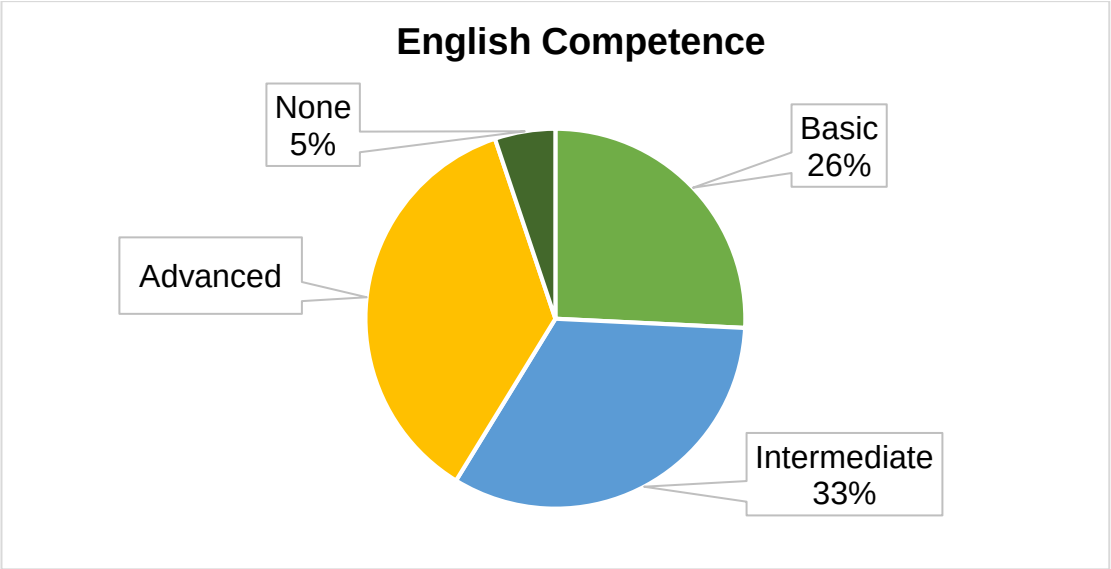

Figure 13. Language competence of respondents

Though there is no entrance requirement for registering as a *Yeeyan* member, the survey result implies that higher education and language competence represent in fact the prerequisites for participating translation-related activities. This corresponds with Shye's (2009) research into the causes of volunteering – the demographic characteristics a person should possess to be able to participate in intellectual work like translation.

The result demonstrates that **people with higher education experience are more likely to participate** in translation-related activities in the crowdsourced environment, believing that language competence confirms L2 learners' ability to translate. It may be a factor that determines the extent of the contribution a person can make to the community. An interrelationship between the language competence and the roles people have played is analysed in section 4.5.1.

4.2 Motivations

Knowing why people are willing to volunteer at the cost of their own efforts and time is vital in understanding the crowdsourcing modes in *Yeeyan*. More importantly, the interest in motivation studies derives to a large extent from a practical concern: knowing why people volunteer can enhance the recruitment and retention of volunteers to human service and other

organisations (Cnaan and Goldberg-Glen, 1991; Gidron, 1985). To fully understand the factors that can influence people's behaviour, the survey explored not only what motivates participation but also what de-motivates participation.

Prior to conducting the survey, my assumption for the motivations that drive people to contribute in *Yeeyan* largely included intrinsic factors, as most of the activities are free or poorly paid. Moreover, according to Participatory Culture, the belief of engaging to make a contribution regardless of the role and effort seems to support the assumption. However, the survey reveals that extrinsic factors also take a vital position in determining whether a person wants to participate or not.

This section analyses the motivations from two broad perspectives: intrinsic reasons (psychological factors), and extrinsic reasons (external benefits). It discusses the findings of the self-selected and ranked motivations that were captured based on Cognitive Evaluation Theory, Self Determination Theory and General Interest Theory. It also reports the self-proposed motivations and de-motivations collected from the open comment area of the survey.

4.2.1 Intrinsic motivations

In this research, I use Ryan and Deci's (2000:56) definition of *intrinsic motivation*: 'the doing of an activity for its **inherent satisfactions** rather than for some separable consequences. When intrinsically motivated, a person is moved to act for fun or challenge entailed rather than because of external prods, pressures, or rewards.' It is related to the three basic human needs in Cognitive Evaluation Theory: the need for autonomy, the need for competence and the need for social relatedness.

Table 7 provides an overview of the respondents' intrinsic reasons for participation, which have been broken down by the types of motivations. As it shows, the expected intrinsic motivations were generally present.

Table 7. Intrinsic reasons for participation (the statements included in the third column in Table 7 and the fourth column in Table 8 are randomly listed in the questionnaire survey)

Reasons			**Percentage**
Intrinsic Motivations	***Altruism***	To make foreign information available to Chinese readers	55%
		To help support the initiative of *Yeeyan*	13%
	Community identification	To network with others and make more friends	23%
	Enjoyment	For fun	50%
		Because the project is intellectually stimulating	20%
		I am interested in the topic of the ST	42%
	Self-achievement	For the sense of self-achievement	40%

Proponents of crowdsourcing development emphasise the selflessness of participants (Lakhani and Wolf, 2005; Olohan, 2014). More than half of the respondents sought to make foreign information available to Chinese readers (55%), which in fact is in line with the initiative of *Yeeyan* (13%). Both reasons suggest a variant of intrinsic motivations – **altruism**, in which one seeks to increase the welfare of others. It is the personal disposition at the opposite pole from selfishness – 'doing something for another at some cost to oneself' (Ozinga, 1999). It is a generally believed reason for all volunteer activities. The survey verified 'altruism' as an intrinsic motivation in the *Yeeyan* community. For example, one respondent stated:

'我爱读日本文学，学习日语之后接触到原版书，当时获得的共鸣和感动比阅读翻译著作所感受到的更甚，希望和我的小伙伴们一起把曾经感动我们的文字翻译成中文，哪怕能够多传递给一位读者也是好的。另外，我还希望告诉读者，日本的优秀作家并不	*'I **like** Japanese literature. I started to read original editions when I learned Japanese. I was moved and felt myself resonate more with the original work than reading the translated version. I wish my friends and I can **share** the words that once touched us with more readers through our translation, even it is just one more reader. Plus, I want to **tell the readers** that there are more*

止川端康成、村上春树或东野圭吾那几位畅销书作家，还有更多作家的作品（由于是距今时间久远的一些作品）闪现着人性的光辉，具有宏大却含蓄的东方之美。’	*excellent Japanese writers besides the well-known best-selling authors such as Kawabata Yasunari, Haruki Murakami and Higashino Keigo. There are more literatures (works date back to antiquity) that spark the brilliance of humanity with implicit but immense Oriental beauty. ’*

Meanwhile, a similar percentage of users stated that they were participating for ‘**enjoyment**’. Respondents agreed on the positive consequence of performing activities via *Yeeyan* – psychological needs can be satisfied. Personal hobbies and preferences may decide the degree of enjoyment from certain behaviours. ‘For fun’ accounts for 50% of the participating reasons. Just like playing games or watching movies, participants who selected this reason naturally like to engage in translation-related activities in *Yeeyan*. Meanwhile, reasons like ‘Because the project is intellectually stimulating’ (20%) and ‘I am interested in the ST’ (42%) indicate the influence of personal preferences in getting people involved. It is in accord with the *General Interest Theory* that emphasises the relevance of the task or, in other words, the level that task context or task content helps satisfy psychological needs and desires.

Another intrinsic motivation, ‘**community identification**’ is derived from ‘altruism’. Psychological researcher Hoffman (1981) defines this type of behaviour as ‘kin-selection altruism’, which corresponds to one of the psychological needs in *Cognitive Evaluation Theory* – the need for social relatedness. Social networking is a commonly acknowledged factor of a majority of the activities performed on the internet. It can fulfil the human needs for belonging. My survey results indicate that participating ‘to socially network and to make friends’ accounts for 23%, which is slightly lower than other reasons. Respondents identified themselves as members of the *Yeeyan* community and aligned their goals with those of the community. They may treat other members of the community as kin and thus be willing to do something that is beneficial for them but probably not for themselves. It may also explain the behaviour of offering feedback and suggestions within the *Yeeyan* community. People see it as a ‘fruit of all’, which is encouraged

by the sense of ‘community identification’. For example, one respondent stated an extra reason:

*‘To **share my experience** as a translator, especially with beginners.’*

The last intrinsic reason, ‘self-achievement’, was acknowledged by 40% of the respondents. It verifies another need in *Cognitive Evaluation Theory* – the need for competence. It contains a process of self-actualisation in which people have the opportunity to confirm one’s self-esteem. More than 1/3 of the respondents who agreed with the reason of sharing foreign information also selected self-achievement as a motivation. A link between altruism and self-achievement is revealed. As respondents put it:

‘我喜欢性学，希望引进先进的性学资讯。通过翻译实现了自己的目标，引入了大量性学资讯。一份翻译就是一份贡献吧，一份支持就是一份贡献吧。’	*‘I am **interested** in sexology. I wish to introduce advanced sexology information to China. I **achieved my goal** through translation and had **introduced a substantial amount of information** about sexology. **Every piece of translation contributes, every activity of support contributes**.’*

‘自我欣赏与自我满足的同时，还娱乐了别人~’	*‘I get to **appreciate myself** with the feeling of **self-achievement** while **entertaining others**.’*

4.2.2 Extrinsic motivations

Extrinsic motivation is defined as ‘a construct that pertains whenever an activity is done in order to attain some **separable outcome**.’ (Ryan and Deci, 2000:60). *Self-determination Theory* proposes that extrinsic motivation can vary greatly in the degree to which it is autonomous: whether the **intentional** act can achieve **instrumental value** (ibid).

Table 8 presents an overview of the extrinsic reasons of participation. Motivations in *Human capital* and *Personal needs* may also be considered as intrinsic reasons, here I follow the studies of volunteers for *FOSS* initiatives (Lakhani and Wolf, 2005) and *Wikipedia* (Dolmaya, 2012). The categorisation

of extrinsic motivations in these two studies are based on the economic theory, in which the net benefit of participation consists of (1) immediate payoffs, e.g. being paid to participate, and gaining the direct use of the resulting product; (2) delayed payoffs, e.g. improved skills, and strengthened prospects of career advancement (Lakhani and Wolf, 2005:6-7). Therefore, the reasons under the category of *Human Capital* are seen as delayed payoffs which lead to external instrumental value of participation; while the reasons under *Personal Needs* are immediate payoffs in which the participants are able to use and benefit from the translation that they are helping with.

Table 8. Extrinsic reasons for participation (the statements included in the third column in Table 7 and the fourth column in Table 8 are randomly listed in the questionnaire survey)

Reasons				**Percentage**
Extrinsic Motivations	***Current and future rewards***	Human capital	To improve language abilities	73%
			To gain translation experience	63%
			To get an early look at translation environment	22%
		Self-marketing	To enhance my reputation as a translator	32%
			To attract translation clients	5%
			To gain the opportunity to have my translation published	17%
		Peer recognition	To gain recognition within the community	13%
			To earn honourable badges	5%
	Personal needs		I need the information to be available in TL	15%
			To voice my opinions on certain topics through my translation	10%

First of all, no respondents selected 'to earn money' as a motivation. Since only participants in the *Gutenberg Project* are rewarded with royalties in *Yeeyan* and 38% of the respondents had participated in it, it seems that monetary reward was not valued by these respondents. Among all the extrinsic motivations, 'to improve language abilities' is the most favoured reason (73%), followed by 'to gain translation experience' (63%), which can all be interpreted as a form of future rewards through the accumulation of human capital. All respondents who are currently students selected these two as primary reasons. It implies that they see the crowdsourced initiative as a training environment, as suggested by O'Hagan (2008) with respect to fan translation networks. It is also in line with the findings of the Dolmaya (2012) study into the crowdsourcing translation initiative of *Wikipedia*. 67% of the respondents to my survey reported that participation in *Yeeyan* helped improve their language and translation ability. At the end of the survey, under the open question of 'Please think about your past experience in *Yeeyan*, what have you gained?', respondents wrote*:*

'Valuable experience in translating; Improved language proficiency, both in English and Chinese; Friends including translators and readers.'

'1.对自己当前的水平有个比较清楚客观的认识；#2.词汇量增加；#3.语感更好了；#4.翻译能力有提高。'	*'1. I had **a clear and objective understanding of my present [translation] ability**. 2. My **vocabulary** has been improved. 3. My **sense of linguistics** is getting better. 4. My **translation ability** has improved.'*
'从大佬（对其它译文）的评论那里学到了很多。（单个单词的精密分析，逻辑性分析等。）'	*'I **learned a lot** from the senior members' comments (on other's translations). (**Precise analysis** of individual word, **logical analysis**, and so on).'*
'其实就是感觉翻译文章带来最大的收获就是可以对自己喜欢的原文有更深的理解，翻译完了以后成就感很大，也增加了我的阅读兴趣吧。'	*'Actually, I feel that my biggest gain from translation is that I had a deeper understanding of the source text that I love. After the translation is finished, I have a strong feeling of self-achievement. It also reinforced*

	my interest in reading.'

'主要是初期接触翻译的时候，实践太少，来译言积累一点实践经验。通过参加古登堡计划，磨练了我的翻译水平和翻译速度，了解了众包翻译的操作过程。'	*'When I first started translation, there were few opportunities for practice. I joined Yeeyan to **gain real-life experience**. Through participating in the Gutenberg Project, **my skill and speed of translation have been improved**. I also **learned the process of crowdsourcing translation**.'*

On the basis of the comments in the open questions, more motivations are extracted. Apart from the benefits of language and translation ability, respondents also indicated that *management* skill is a goal that they seek for, mainly for project management and time management. They are all counted as a reward for one's capability – 'human capital'.

Economists use the term 'human capital' to designate personal skills, capabilities, and knowledge. Increasing one's human capital by means of education, training, learning and practising leads to better job opportunities and more fulfilling jobs (Becker, 1962). The properties of *Yeeyan* as a crowdsourcing initiative encourage participants with this motivation, especially for entry-level translators, such as college students and amateurs who wish to enter the professional industry. Extrinsic reasons for enriching human capital and the respondents' self-declared improvements in translation and other skills demonstrate the potential of being a training environment in crowdsourcing initiatives like *Yeeyan*.

Self-marketing is another reason that forms the future reward for participants. 32% of the respondents admitted the motivation 'to enhance my reputation as a translator', followed by 'to gain the opportunity to have my translation published' (17%). Translation products of *Yeeyan* are usually promoted via popular social networking sites in China, such as *Weibo* (a microblog, similar to *Twitter*) and *Wechat* (an instant messaging tool that functions the same as any other social networking site), which makes the translation reach a large audience. It is a cost-friendly platform for freelancers or trainee translators to advertise themselves. For example, one participant

reported that he/she gained:

> *'Some reputation as a qualified translator so that some publishers have approached me for cooperation. I did translate two books myself afterwards.'*

Yeeyan is seen as a good advertising channel for those seeking to advance in the Language Services Industry. Self-marketing here has an important implication. The larger a member's contribution to the community, the more likely that commercial clients will recognise his/her value, thereby increasing the incentive to apply these skills in a paid job. It is empowered by the openness of the crowdsourced environment, in which translators gain visibility from different perspectives. For example, the number of subscribers indicates the extent to which one has been recognised in the community, the number of translations indicates the efforts one has devoted, the type of translations indicates the field that one is interested in or good at. It facilitates the transfer from volunteer work to profitable commercial work for novice and amateur translators.

Peer recognition is another external gain that is associated with extrinsic rewards. It is derived from the desire for fame and esteem (Maslow, 1970). A member who completed a certain amount of translation, worked as a project manager, participated in the *Gutenberg Project* or has been listed on the 'Featured' section in the front page of *Yeeyan* can earn the honourable badges. It is a transparent incentive mechanism that highlights one's contribution which is also associated with peer-recognition. The overall proportion of respondents who were motivated by peer recognition is quite small ('to gain recognition within the community' – 13%; 'to earn honourable badges' – 5%). However, nearly half (47%) of the respondents indicated that they have followed one or several 'famous' translators in *Yeeyan*. And 62% of the 'famous' names they proposed overlap with each other's statement. It shows that peer recognition can be achieved through one's performance in the crowdsourced environment.

'**Personal needs'** is a strong motivator for the development of open source software (Hars and Ou, 2002). It has been found that open-source projects

are often initiated because a programmer has a personal need for a certain kind of software. Similarly, in *Yeeyan*, people also participated to make the information that they need available (15%). They achieved this by providing the desired source text to members, organising the translation task, or directly translating, as well as revising the desired translation. The amount of interest groups available in *Yeeyan* supports this argument. For example, there are interest groups in technology, business, literature, health, nature and culture. Members can join these groups to fulfil their personal needs in various ways. For instance, there is an interest group in Vaccinology: the group members are all postgraduate research students in the subject domain. They gathered to translate academic materials so that they access the latest developments in the field. Moreover, those whose language skill is not sufficient for translation-related activities can always suggest the source text to other members. They usually play the role of loyal audience in the community.

A small proportion (10%) of respondents indicated that part of their reason is to voice their opinions through translation. This is a new finding in motivation research in crowdsourcing or volunteer translation. Respondents with this reason usually decided whether to participate or not according to the topic of the source text. Other reasons linked with personal needs were proposed:

> *'I want to get out of those walls and see what the world is actually like outside.'*

> *'feel the different culture, politics, people and other extraordinary things'*

> *'My project on Yeeyan is for my daughter. I contribute my efforts to make it a gift for my little angel. I'm looking forward to her reactions about this project.'*

There was one particular respondent indicating that she translated a book via the *Gutenberg Project* on *Yeeyan*. She replied to my invitation message to the survey and introduced the book she translated. It was a biography on feminism. With the aim to set a role model for her daughter, she hoped that her daughter can gain some insights from the book.

4.2.3 Overall result of motivations

The above discussion explains the reasons for participating in crowdsourcing translation based on the types of motivations. In the survey, respondents were also asked to rank the top 5 reasons in the order of importance. Figure 14 illustrates the ranking result and Table 9 shows the average rank of the motivations.

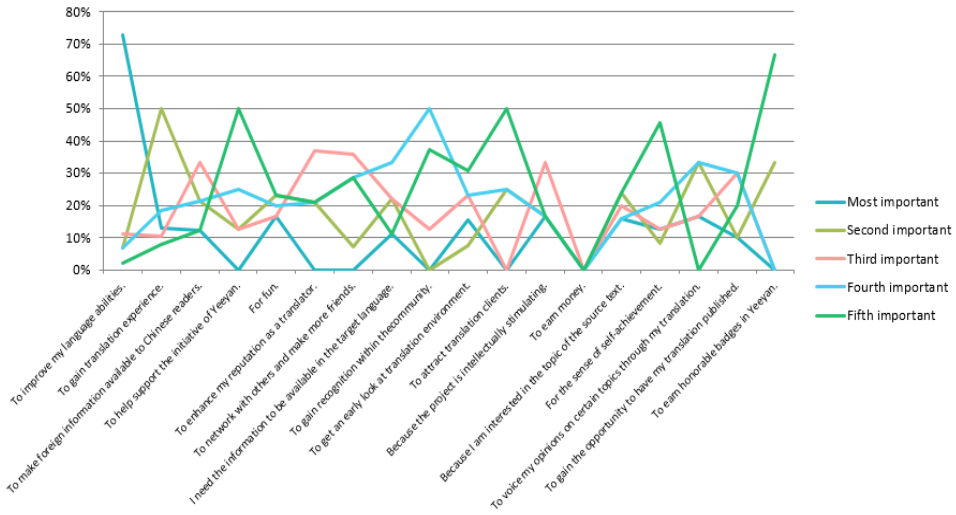

Figure 14. Rank of motivations in the order of importance

Table 9. Average rank [18]of the motivations in the order of importance

Reasons	**Average rank**
To improve my language abilities.	1.59
To gain translation experience.	2.58
To voice my opinions on certain topics through my translation.	2.67
To make foreign information available to Chinese readers.	3.00
For fun.	3.10

A surprising result stands out: according to the importance of reasons,

[18] The formula for the average rank score is $(X_1W_1 + X_2W_2 + X_3W_3 \ldots X_nW_n)$/Total, where W = weight of the ranked position and X = response count for answer choice.

extrinsic motivations exceed intrinsic motivations. Future rewards and personal needs seem to be more attractive than participation for altruism and enjoyment. It corresponds with Hars and Ou's (2002) study in which a correlation analysis is used to examine the significance of motivations. They concluded that external factors are more important than the internal motivations in the open source environment. The crowdsourcing community in the case of *Yeeyan* resembles the open source environment, in which product-providers are also the product-consumers. The identity of participants and the properties of these two environments overlap to some extent, which explains the similar result in motivations.

Given that only 10% of respondents chose ‘to voice my opinions on certain topics through my translation’ in the first place, it is necessary to explain why it occupies the third most important position in the average ranking of the motivations. Statistically speaking, it was because these who selected the reason ranked it as the ‘second’ or ‘third’ important motivation. One common feature between these respondents was that they all chose whether or not to translate an article depending on the ‘topic of the source text’. Yet, their translation strategy was either in line with the project brief or stayed close to the source text. Only one respondent indicated he/she usually translated and edited the text as they wished. It seems that ‘voicing one’s opinion’ is more likely to be achieved through what people translate (i.e. the source content about which they share similar opinions), rather than through the way in which they translate in *Yeeyan*.

As mentioned in section 3.1.2.1 in Methodology Chapter, there is a common concern in survey research in general and in motivation research, in particular – the problem of social desirability. The problem lies in the tendency of respondents to intentionally report motivations that would make them look better (Shye, 2009). In light of this concern, the strategy was to include a question which asks the respondents to select **the potential motivations for which one thinks other people participate** in *Yeeyan*. The results are similar to the previous discussion. Extrinsic reasons take a much larger proportion than intrinsic reasons: ‘To improve language abilities’ remains the top potential reason (83%), followed by reasons as: to gain

translation experience (80%), to make foreign information available to Chinese readers (62%), for the sense of self-achievement (53%), and to gain the opportunity to have one's translation published (52%).

At this stage, the study confirms that: first of all, people are encouraged to volunteer in *Yeeyan* for multiple reasons (both intrinsic and extrinsic); secondly, extrinsic motivations are perceived as more important than intrinsic motivations from the standpoint of *Yeeyan* members.

4.3 Motivations and Behaviours

Studies have investigated the interrelationship between motivations and individual performance, showing that intrinsically-motivated goals (i.e., autonomous) are associated with most effortful behaviours, as compared to controlled personal goals (i.e. non-intrinsically motivated goals), and thus lead to a greater possibility of goal attainment) (Sheldon and Elliot, 1998). In other words: participants with intrinsic motivations will spend more time and effort in volunteer work; people with personal needs have a limited amount of effort that they are willing to provide as volunteers. Is this the same case as in *Yeeyan*?

To find out the interrelationship between motivations and individual performance, I associated the reasons for volunteering with the time that one spent working on translation-related activities in *Yeeyan* per week (self-declared working hours). I first measured the percentage of top important motivations among the respondents. Again, this shows that extrinsic reasons (human capital, personal needs and self-marketing) exceed intrinsic reasons (altruism, enjoyment and self-achievement). The result remains almost the same as the rank of motivations in the previous section. I then calculated the average working hours per week for respondents with the same top important motivation.

Table 10. Percentage of the most important motivations and the average working hours per week.

	Most important motivation	**Percentage**	**Average working hours per week**
Extrinsic	Human capital	68.6%	10.9
	Personal needs	3.4%	12.0
	Self-marketing	1.7%	5.0
Intrinsic	Altruism	6.7%	4.5
	Enjoyment	15%	4.4
	Self-achievement	3.3%	3.0

57% of the respondents chose ‘to improve language abilities’ as the top important reason for participation. Under this reason, the longest working time was 72 hours per week, while the shortest was one hour. For people working for personal needs, the working time ranged from 1-23 hours. Enjoyment appears to be the top reason among the intrinsic motivations with a time range between 1-15 hours. The working hours of those who contributed primarily for altruism and self-achievement were: 1-10 hours and 2-4 hours per week, respectively.

Compared with Sheldon and Elliot's (1998) research in social psychology studies, it seems that a contrary result has appeared. However, people’s behaviours are usually influenced by multiple factors: controlled and non-controlled factors. At this stage, I consider only the apparent relationship between two variables (i.e. the most important motivation and the time investment in translation-related activities in *Yeeyan*). There should be more influential factors that decide the participants’ behaviours. As the respondents stated:

‘It depends: sometimes none, sometimes up to 20 [hours]. ‘

‘8 小时，不好说，取决于参加的项目情况，时长时短。’	*‘Eight hours maybe. Hard to say. It depends on the settings of the project that I participate in; it can be longer or shorter.’*

The following part explains other self-declared factors that may influence people's performance in the crowdsourced environment.

4.4 De-motivations: what are the possible influential factors?

By asking 'from your past experience, what may stop you actively participating in *Yeeyan*?', various de-motivations arise.

4.4.1 Settings and properties of the project

To begin with, respondents indicated that the settings and properties of the project can influence their performance, the **source text** in particular. According to the survey, 88% of respondents chose whether to translate an article or not depending on the topic of the source text. This corresponds to the intrinsic motivation of enjoyment. Personal preference seems to be vital in determining which activity to perform, with 28% of the respondents valuing the source of article. For example, there was one respondent who was interested in technology who indicated that he/she focused on '*articles from certain websites, such as The Verge*[19].' Another possible reason may be that the participants feel more confident in understanding the source text which they are interested in or of which they have specialised knowledge.

Other distinct factors are the length and degree of difficulty of the source text. If it is a relatively long text, they would prefer to collaborate with others than work alone. With regard to the degree of difficulty, during my observation I found that specialised texts, such as health texts, can push away the novices in the community even though they committed to the project of their own accord. In the questionnaire survey, 30% of the respondents indicated that 'source text is too difficult' can de-motivate their participation and may lead them to quit in the middle of the project. Chapter 6 points out how withdrawal or delay of the project can be noticed in advance through *Dialogue Act*

[19] 'The Verge' is an American technology news and media network which offers in-depth reporting, long-form feature stories, and technological product reviews. Available from https://www.theverge.com/.

Analysis and a suggestion for solving the problem is proposed in Conclusion.

In terms of the settings of the project, two influential factors emerge: **schedule** and **rules**. Scheduling is a crucial aspect of project management. It closely associates with the flexible design of the project: the workflow and up-to-date progress. In *Yeeyan*, a collaborative project can be organised by the one who provides the source text, the Editor (a recruited position in *Yeeyan*), or one volunteer interested in project management. They are considered as the project manager who makes schedules and provides guidelines. During the observation, I found that some of the projects had gone beyond the agreed deadline. The frequent case is that the preparation and translation stages progress well, but the revision stage tends to take much longer, which leads to a delay in final delivery. Chapter 5 (section 5.2.1) contains a detailed discussion of translation and revision productivity in the observed projects with reference to the productivity in the professional industry and also when working with translator trainees. In the survey, respondents indicated that they decided whether to participate based on the length of the project. 52% of the respondents stated that if the process takes too much time, they may end up quitting or not signing up for the project.

Moreover, 23% of the respondents indicated that ‘too many rules to follow’ may also de-motivate their performance. Here, rules refer to the requirements indicated in the ‘project brief’ as in the professional industry. It mainly includes the desired tasks, deadlines, terminology standards, usage of tools and sometimes style guidelines. As mentioned in Literature Review chapter, ‘autonomy’ is one of the core principles in Participatory Culture which lays the foundation of crowdsourced environment compared with the professional industry. Users need a sense of autonomy during participation. Self-governance is emphasised in such an environment. Redundant rules can decrease enjoyment because many people perform activities in their spare time. After a full-day work or study, one naturally does not want to engage in an environment that does not allow freedom. However, it does not mean that people in the crowdsourced environment do not have the ethics to properly fulfil one’s task. The following chapters further discuss this concern with reference to the projects I observed.

The final factor lies in the **competence of team members**. Potential participants can measure the ability of the group as a whole before joining a project. It is the result of a crowdsourced environment in which the status of an individual can be seen. Moreover, 15% of the respondents indicated that their decision on whether to contribute in a collaborative project depends on who they are going to work with. It is also the consequence of peer recognition in community culture. 27% of respondents stated that the low quality of group members' performance can de-motivate their participation. To examine the validity of human reason as an influential factor, the survey asked respondents to name 'famous' translators that they know or subscribe to in the *Yeeyan* community. As mentioned in section 4.2.2, nearly half of the respondents followed at least one high-level translator. I checked the profile of these 'famous' translators and found that they are all experienced members who have a substantial number of translations and earned several honourable badges. This confirms the feasibility of measuring the competence of team members before joining a project, and also the influence which human factors have in a crowdsourced environment on getting people involved. The respondents' concern regarding the translation competence of team members can also be interpreted from two perspectives: the pursuit of 'good' translation resulting from a collective effort; and the demand for effective peer feedback, which is expected by all contributors in collaborative projects.

4.4.2 Communication and feedback mechanism

Another concern relates to the human element in a project. Respondents demonstrated that the atmosphere within the team matters. **Effective communication** is often expected between team members even though it is not compulsory in the community. As discussed in section 4.2.1, one of the intrinsic motivations regards social networking to fulfil the need for belonging and social relatedness. If a person joins a crowdsourcing project particularly for this reason but ends up with a team that does not actively interact with each other, it is natural for the person to withdraw.

More importantly, since a majority of people wish to improve their language

and translation abilities through participation, communication is needed to support the motivation because it is a convenient way of giving and receiving feedback in the crowdsourced environment. Raymond's (2001) study on open-source software emphasised that the success of open-source software is to a great extent due to its early, fast and frequent releases. As a consequence, open source programmers receive rapid and constructive feedback about the quality of their work. Feedback always has a positive effect as it shows programmers that people are using their contributions. The argument is also supported by Hars and Ou's (2002) research on the motivations for participating in open-source projects. Feedback mechanisms in crowdsourced environments like *Yeeyan* are an essential component that fosters the thriving community.

Peer interaction and translator-reader interaction offers direct feedback in terms of translation quality and helps achieve peer recognition. Such interaction shows that people's work is valued, which in return improves one's translation competence. 100% of the respondents who selected 'to gain recognition within the community' also chose 'to improve language abilities' as their primary reason for participating. Thus, the feedback mechanism is self-reinforcing, for it encourages participants to dedicate additional effort to perfect their work, which in turn attracts more *favourable* feedback.

The need for communication is also manifested in the end of the survey when asking for people's suggestions about *Yeeyan*. 1/6 of the respondents expressed the same opinion that more communication is needed and hope that *Yeeyan* establish some mechanisms that further encourage the interaction between members either to discuss translation problems or to social networking. 34% of the respondents demonstrated that if it is too hard to communicate with team members, they may withdraw from the project. 57% of the respondents showed that they were particularly in favour of the communication aspect in collaborative translation because it can help improve the work quality and bond with other members. Professional industry has already seen the importance of communication in collaborative translation projects. *MemoQ 2014 R2* is the first CAT tool that integrated an

instant messaging function. The discussion feature is not only a transparent way to communicate, but also helps the translation company build a knowledge base which is an important resource that shortens the learning curve of translators (Kilgray, N.D.). Chapter 6 discusses the real circumstances of peer interaction in the crowdsourced environment and investigates how communication influences performance using data collected from my observations.

4.4.3 Personal factors

Personal factors refer to the 'distractions' that may negatively influence the participants' involvement in the community. Respondents declared that they would stop contributing to *Yeeyan* if they were too busy with their main work. Moreover, as mentioned when discussing extrinsic motivations (section 4.2.2), one of the benefits in *Yeeyan* is that people get to transfer volunteer work to profitable commercial work through the self-advertising effect. One respondent stated:

> '*Since I can now easily get paid for translating similar content, I choose not to translate for free in Yeeyan*.'

Three respondents stated that they reduced the time they spent in *Yeeyan* when they began to receive paid translation/interpreting tasks. Hence, it is reasonable to treat a crowdsourced environment as a good practice-ground for novice translators and amateurs to experiment in. This suggests a tendency that when the initial goals have been fulfilled, such as improving language abilities and self-marketing, one may devote more effort in the professional industry rather than volunteering through crowdsourcing.

Yet, there are still a great many of the long-term members who vibrantly participate in various activities in *Yeeyan*. This happens because, on the one hand, they are well-recognised in the community and, on the other hand, *Yeeyan* launched a commercial service in 2009 which offered the opportunity to make a profit by offering high-quality required activities, such as the *Gutenberg Project* and other paid services. It is a flexible approach to maintain the user-base, which highlights the importance of understanding the

crowd's motivations. Once the organiser knows why people are participating, they can adjust the operation modes to fulfil these intrinsic and extrinsic goals in a way that benefit both the user and the community.

Additional personal factors including health issues or personal reasons that may also influence the vitality of members. Such factors cannot be easily controlled, but can still be managed by a competent project manager. A lead user or project manager needs to be alert to such unforeseen circumstances and be ready to make flexible plans (further discussion see Chapter 6).

4.4.4 Conditions of the crowdsourced environment

In the survey, several conditions of the crowdsourced environment were proposed to be relevant in deciding whether or not to participate in translation-related activities in *Yeeyan*. Respondents demonstrated concerns regarding: 1. the **user-friendliness of the platform**; 2. **the evaluation metric of user performance**. Respondents expressed dissatisfaction towards the text-editor built for translation in *Yeeyan*. Moreover, the 'headnote' function for making suggestions on other people's translations was sometimes difficult to use, as well. They also indicated that the homepage was no longer user-friendly after the several rounds of webpage and structure update. Though the community is divided based on interest groups, it is difficult to find a preferred source text in the defined category. The way in which source texts are presented and organised seems messy. The storage and the quality of source texts is not sufficient for members with high requirements:

'译言网的原文库储量不是很够而且质量不是很高啊==，有时候经常会遇到相当杂乱的文章# 当年人气鼎盛的好时光，真的已经过去了'	*'Yeeyan does not have sufficient **source text storage** and the **quality** is not high. Sometimes I came across very **messy** articles in the source text database. The good times when Yeeyan is of great prosperity has gone.'*

'我认为，文章应该更细化一些，例如把文章按照难度分等级，让译者根	*'I think source texts can be classified with more detail, for example, **classify the source texts***

据自己的水平来选择原文。'	***according to the degree of difficulty**. In this way, we can choose the articles that match our ability.'*

On the homepage of *Yeeyan*, there is a 'featured' section (译言精选) to publicise good translations which are chosen by the editors. It is seen as an evaluation of user performance in the community and a straightforward approach to gain peer recognition, as well as to advertise high-level translators. Respondents demonstrated that the evaluation standard of the 'featured' articles is decreasing. Fewer high-quality translations are listed. For example, respondents commented:

*'**Poor quality** in translation products that are chosen to be 精选 (Featured).'*

*'The **quality** (and frequency of publication) of the 'Featured' articles appeared to have come down quite significantly in the past one year or so. I **used to enjoy reading** that portion of the site on a daily basis – bilingual or translated text only. Now I can **hardly** find a good (translated) article to read in a week. Sometimes the quality is so bad I would just go directly to the English version for the read. #Hope Yeeyan puts its quality hat back on, and I am sure more translators will join / come back to the community.'*

*'They need to work on the (译言精选) **featured** selection process. It's really reflecting Yeeyan very poorly.'*

Overall, the respondents' feedback implied that **a fair, transparent and high-level quality evaluation mechanism is needed** both to positively motivate participants and to maintain the user-base. They also expected the editing platform to feature user-friendly collaboration tools for translation and revision.

4.4.5 Non-controlled factor – censorship in China

Censorship was brought up in the survey. Though I aim to investigate only the current situation of crowdsourcing translation in China and to avoid any problem relating to the political system, respondents demonstrated concerns about censorship in content-generation. It was proposed in the final open-

question which asked for suggestions and expectations that members have for *Yeeyan* or for the participants themselves. Four respondents indicated that the strict censorship sometimes de-motivates their participation. For example, one stated:

> *'It [Yeeyan] has been suffering from censorship too much, but I don't suppose it can be improved, given the circumstances in China. However, I do hope Yeeyan could make itself known to more people.'*

Censorship has two meanings. One refers to the information control policy in China, meaning that articles with sensitive words or topics should not be introduced to Chinese audiences by the means of translation. In line with the government regulations, employed editors in *Yeeyan* have to filter the source texts that are proposed by members and the translation produced in the community to avoid political problems. Censorship particularly focuses on media reports from foreign countries which is a popular topic that *Yeeyan* members are interested in. Therefore, it is seen as a non-controlled factor that impedes people's participation. Since its foundation in 2006, *Yeeyan* has been requested twice to rectify its website in accordance with the government regulation. The first round of rectification was in the end of 2009, during which *Yeeyan* was shut down for 39 days (Yang, 2010). The latest one was in June 2016 (after I finished the data collection) and it was requested to modify its webpage and user system (Yeeyan, 2016).

Censorship also refers to the internal filter that protects the copyright of source text authors. Text editors need to make sure there is no infringement of copyright in terms of the content generated on *Yeeyan*. The consequence of censorship is that members may not have the full autonomy in deciding what to translate. Under the big picture of Chinese politics, crowdsourcing translation in China seems to be not completely user-driven. It can be an influential factor for those who value the autonomy and freedom generally associated with volunteer activities.

4.5 General Participation

To investigate the real practices of crowdsourcing translation in *Yeeyan*, the survey explored what kind of activities were performed in the community, i.e. what kind of roles the participants play to support the ultimate translation practice. It also investigated how the activities were performed, i.e. methods that were used to control the translation quality.

4.5.1 Multiple Roles

The survey indicates that *Yeeyan* members usually play multiple roles in the community. Under the question 'What kind of roles have you fulfilled in *Yeeyan*', unsurprisingly, 'translator' appears to be the most common role in the community (83%). Other roles are: reader (38%), reviser (28%), source text provider (27%), commentator (17%), project manager (17%), promoter (6%), and mentor (2%).

Figure 15 illustrates the dynamics of a fully-functioning crowdsourcing community based on the roles people have fulfilled. The definitions of *translator*, *reader*, *project manager* and *reviser* are the same as in the professional industry. *Commentator* refers to those who like to comment on others' translation or the content of the text. It differs from reviser in that the commentator does not directly engage in the revision process but merely offers suggestions after the translation is published. It is an exclusive form of translator-reader interaction in which post-feedback is provided. The suggestions are open to all readers. Translators can choose whether to implement the suggestions or further discuss them with the commentator. *Promoter* refers to those who share translations from *Yeeyan* to other sites. *Mentor* refers to those who advise others on translation strategies or tools that they could use to improve their translation skills; mentors also answer questions regarding the translation practice.

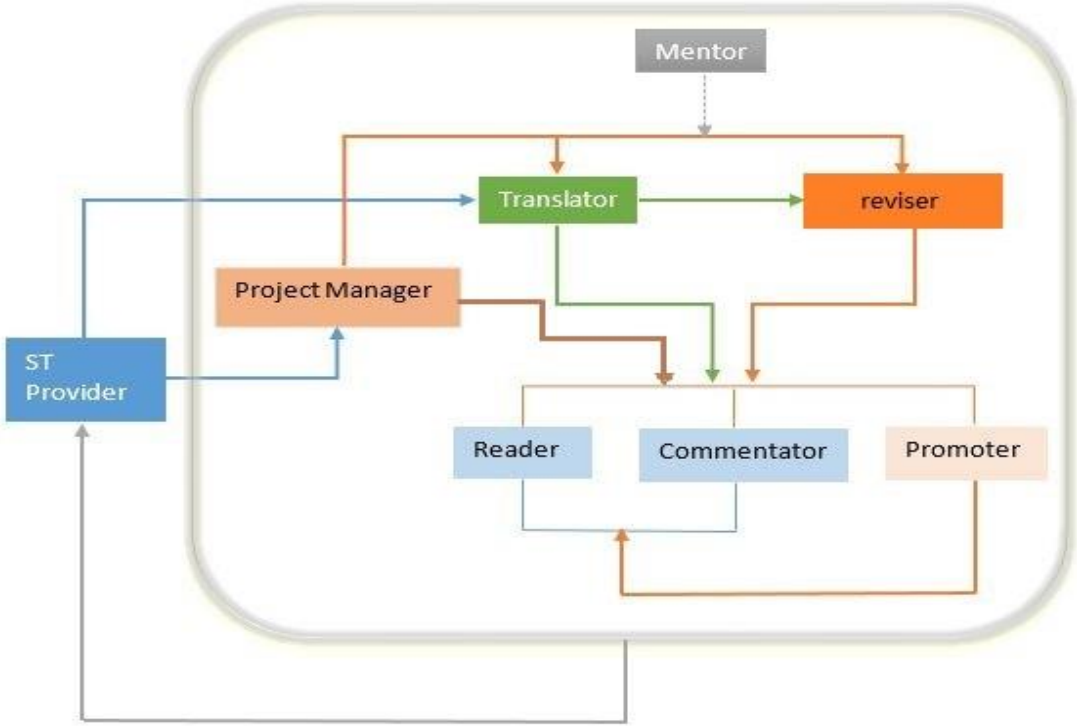

Figure 15. Roles that support the functioning of the *Yeeyan* community

Typically, one individual can play multiple roles during the content circulation and generation process, and it is not necessary for every role to be involved in the process. Independent work and collaborative work both run in the community, which does not usually operate in a top-down model. It is more likely to be a self-organising community in which members spontaneously play the needed roles. Figure 15 merely gives an overview of how members play their roles and collaborate as a community in *Yeeyan*. The evaluation of the efficiency and translation quality in such a dynamic environment is further performed based on the data collected from observation in 6.4.

A statistical analysis was conducted to explore the relationship between participants' qualifications and the roles they have played in the *Yeeyan* community. Since translator, reviser and project manager are the three roles that are directly involved in the translation production process, they are considered as the significant roles in the *Yeeyan* community. Figure 16 demonstrates a simple linear regression between the roles and the respondents' English competence. The *X* axis represents the independent

variables, i.e. the English level of participants, and the *Y* axis represents the dependant variables, i.e. the number of participants under each role. A positive trend was found: **the higher the participants' English level was, the more likely they were to take on significant roles**, i.e. PM, reviser, and translator.

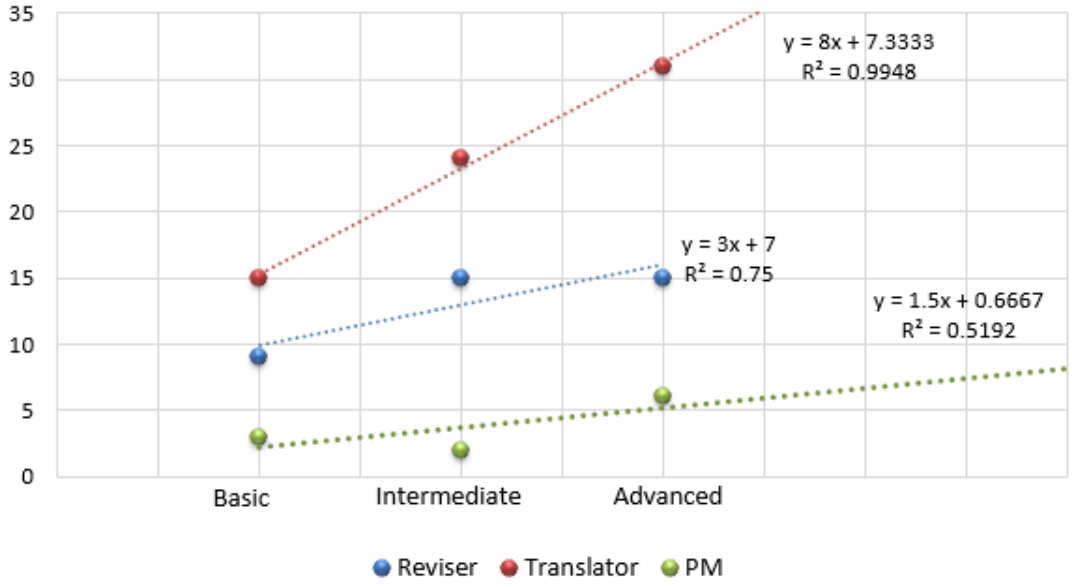

Figure 16. Relationship between English competence and significant roles

Cámara (2015) identified that the reviewing and management positions in the *TED Open Translation Project* were mainly taken by professional translators. I also attempted to investigate the relationship between the significant roles and the experience of the respondents (i.e. years they have been working in *Yeeyan*, education and translation-training background, number of translations they have produced). A similar result arose: **the percentage of respondents whose occupation was professional translators and who did language-related work made a difference, as *revisers* were more likely to belong to these categories than to students**. The percentage of revisers whose occupation belongs to the aforementioned three categories are: 25%, 7%, and 5%. However, no explicit relationship between the significant roles and other parameters related to experience was found.

The following sections discuss the participants' approaches to translating according to respondents' own experience in order to draw a full picture of

the current conditions of crowdsourcing translation in *Yeeyan*.

4.5.2 Perceptions on crowdsourcing: sense of responsibility and ethics

As mentioned in section 2.2.4 in the Literature Review chapter, critics of crowdsourcing translation often focus on the ability of participants to produce quality work, arguing that amateurs do not follow professional ethics codes, cannot solve translation problems properly, and most importantly, the translation quality cannot be assured. This section looks at the respondents' perceptions of crowdsourcing and explores whether *Yeeyan* is confronted with the same problems.

First of all, in terms of the competence of translators, as discussed in section 4.1.1, a majority of *Yeeyan* members have at least the entry-level language ability to perform translation activity as 'natural translators'. Having said that, can they deal with translation problems? According to the survey, only 6.67% of the respondents indicated that they would ignore the problem that they come across during translation and simply do nothing about it. The remaining 92%'s problem-solving strategies included 'consulting subject-matter professionals' (35%), 'researching the issue on one's own' (85%), and 'communicating with other translators/team members' (62%). The results of the problem-solving activity are largely unknown, but nevertheless suggest a tendency for participants to hold a positive attitude towards problems instead of avoiding them. Once again, the high proportion of problem-solving through peer-to-peer interactions highlights the importance of communication and feedback mechanisms in the crowdsourced environment.

Another general view is that since participants are not legally committed to the translation task in the crowdsourced environment, they may not have the sense of responsibility to finish the translation in accordance with the task requirements; for example: one may withdraw from the task out of the blue or miss the deadline (Désilets, 2007); or participants may not have the awareness of quality control, and may treat the task perfunctorily and use machine translation instead of human effort. The assumptions are all possible considering the properties of crowdsourcing. However, the survey

found that a majority of respondents in *Yeeyan* performed the tasks in a similar manner as in the professional industry.

To begin with the translation strategy, 58% of the respondents valued the equivalence between source and target texts, and indicated that they tried to stay as close as possible to the source text during the translation. 23% claimed that they translated according to the project brief, while only 13% stated that they translated and edited the text based on their own will. For example, respondents wrote:

*'I usually **keep the original meanings** and make the reader enjoy the reading.'*

'视原文而定，个人情感为重的文章(如小说、博文)尽可能与原文一致，宣传类的文章（如科普）会稍作编译。'	*'It depends on the **source text** (ST). If the ST is an article emphasising personal attitudes and emotions (such as a novel or a blog article), I try to **stay as close as possible to the ST**; if it is an article for publication (such as popular science), I **may trans-edit** the text. '*

'首先是忠实原文，然后再考虑阅读体验，根据原文风格进行润色。'	*'My priority is to be **faithful** to the ST, then consider the reading experience, polishing it according to the **style** of ST.'*

These self-claimed ideas are basically the same as how translations are processed in the professional industry. The majority of respondents demonstrated their awareness of translation ethics. No malicious translation act was revealed in the survey. Plus, 57% of the respondents indicated that they read through the project brief when they were involved in collaborative projects; half of them always followed the instructions and requirements, while the other half would communicate with the PM if there was anything in the brief they did not understand, or they thought was negotiable. This is also reflected in the observed projects which are analysed in Chapter 6. Those who value the project brief were primarily motivated by extrinsic reasons (human capital – 74%).

As for the concern of treating translation perfunctorily and use machine translation (MT) without proper post-editing, only 10% of the respondents indicated that they had such experience. While 30% of the respondents declared that they never used MT output during their translation production in *Yeeyan*. 8% of the respondents demonstrated that they post-edited the MT to improve quality before actually using it, and 52% usually used MT output as a reference to lookup vocabulary, specialised terminology and expressions.

In terms of the sense of responsibility, 55% of the respondents showed that they would give team members or the PM notice if something unexpected occurred that forced them to withdraw from the project or could cause the delay of their work. 10% expressed that they would not take action when such circumstances occurred. Under the question of 'have you ever disappeared/quit in the middle of a project before completing the task', 15% of the respondents acknowledged having such an experience. Very importantly for future implementations of this research into professional practice, the observed projects demonstrated that a participant's withdrawal can be anticipated from his/her performance in the team communication (discussed in Chapter 6). The Conclusion (Chapter 8) contains suggestions for addressing this issue.

Moreover, I found that 70% of participants who had never quit in the middle of a collaborative project were primarily motivated by extrinsic reasons (human capital – 64%, peer recognition – 6%), while the remaining 30% were motivated by intrinsic reasons (enjoyment).

Overall, through the statistics from the survey, it is reasonable to argue that a majority of participants in *Yeeyan* have the same sense of responsibility and ethics as the ones in the professional industry, especially those with strong extrinsic motivations for contributing.

4.5.3 Quality control

Both in the professional industry and in the crowdsourced environment, the quality of the translation is the ultimate criterion. From the perspective of participants, this section explores how quality is controlled in crowdsourcing, and whether quality control methods differ from the professional industry.

Translations done for knowledge-sharing do not require as high a level of quality as required by the industry, especially given the fact that a large proportion of the members in *Yeeyan* participate to improve their language abilities. Gist-level quality is actually acceptable to the crowdsourcing audience. The aforementioned 'featured' listing is one quality indicator from the perspective of initiator in *Yeeyan*. However, as some of the respondents complained about the quality of the 'featured' translation, simply taking it as the only measure may not be a reliable method of analysis. The common quality control mechanism in crowdsourcing follows the following pattern: after the translation is finished, other members can review the work through a comparative view of the source text and target text to pick out flaws or use headnotes to offer suggestions. Quality control in this mode lies on the combined shoulders of translators, revisers and commentators.

When it comes to needing translations for publication, the preselection of contributors is a frequently-used approach to ensure quality. In the survey, 52% of respondents had taken a sample test before they were admitted to such a project, while 17% demonstrated that their past experience participating was merely to sign in and start translating straightaway. The benefits of sample tests are: on the one hand, the competence of the participants can be examined, and the less qualified translators can be filtered; on the other hand, since having several translators collaborate on the same text in a project is still questioned for reasons of coherence and consistency (Yahaya, 2008), the organiser can recruit participants with a similar translation style through the sample test (indicated by the PM in the observed *Gutenberg Project*). This matches the approach of the European Commission translation services (Vlachopoulos, 2009), where the first stage of recruitment and preparation precedes translation itself with the specific

goal of enhancing the quality of the service. This highlights the importance of recruiting the best translators on the market – 'lead-users' in the words of Howe (2008); while the best translators may not always be available for a crowdsourcing project, using translation tests still gives the option to select the better ones among those who are willing to contribute.

The standard professional TEP approach is also used in 'translation for publication' projects, such as the *Gutenberg Project*. After selecting the appropriate translators and project manager, members work in a private environment which others cannot access. According to the guidelines provided by *Yeeyan*, the usual workflow is to translate, then perform a comparative revision, and finally a monolingual review. Professional editors are involved as gatekeepers in the final stage to assure the quality of 'for-publication' projects. They decide whether the translation is good enough for its purpose and what kind of revisions should be implemented.

During my observations, I noticed that the last two stages of the TEP workflow can be chaotic in the sense that repeated rounds of revision (both bilingual and monolingual) occur. This can be time-consuming and inefficient, yet demonstrates that participants are willing to make the effort to enhance the quality. The assessment of the workflow and productivity is discussed in Chapter 5 (section 5.2.1).

According to the survey, the methods which the respondents had used to ensure quality and keep consistency are:

- Using a shared glossary list (38%).
- Frequent discussion with team members (35%).
- A subject-matter professional or project manager to do the final review (25%).
- Peer-revision (85%).

The result demonstrates a diversity of quality control methods in *Yeeyan*. Given that more than half of the respondents mentioned their awareness of following project requirements and professional ethics, a top-down quality control approach may be feasible in the crowdsourced environment.

However, a significant challenge is that the top-down approach is the ideal model for a well-structured organisation (similar to the strict project management in 'for-publication' projects). When it comes to the general translation projects mainly aimed at knowledge-sharing, my observations of the four *General Collaborative Projects* indicated that the workflow did not follow a sequential order as in TEP. Instead, translation and revision activities were performed simultaneously, in which self-revision and peer-revision were combined randomly. To examine the effectiveness of the participants' self-claimed quality control methods, I present the results of the error-based assessment in 6.4.

Apart from translation, the survey also explored how revision is performed according to the respondents' preference and experience.

4.5.3.1 Self-revision

95% of the respondents indicated that they usually perform self-revision, among whom 58% preferred to self-revise immediately after the first draft of the translation is finished, 37% liked to leave the work a few days and then read it again to see if there is anything needs to be improved. Moreover, 93% of respondents had revisited their work to check others' comments and suggestions after the translation was published on the site. Table 11 illustrates the percentage of errors that respondents stated that they normally corrected during self-revision.

Table 11. Frequent errors corrected during self-revision

Errors Types	Percentage
General vocabulary	68%
Specialised terminology	33%
Grammar	35%
Punctuation	22%
Misspelled character	43%
Style and register	42%
Inconsistency within the text	48%

Based on their answers, respondents had usually corrected content errors present in the glossary, and fixing style-related problems (inconsistency and

register), as well as hygiene ones (misspelled character). On the other hand, linguistic errors (grammar) and specialised terminology were corrected less often (35% and 33%, respectively). Punctuation was the least corrected error during self-revision (22%). These self-declared results are further compared in 6.4 with the empirical data collected from my observations. Interestingly, though nearly all respondents declared that they conducted self-revision in the production process, error annotation of the empirical data suggested that there was still a large proportion of errors left uncorrected through their own efforts.

4.5.3.2 Peer-revision

Considering the respondents' perceptions of peer-revision, the survey explored how participants value other members' comments and revision. 63% of the respondents found peer-revision and others' comments useful in improving quality, and only 8% of them held the opinion that 'when my translation is done, my job is done' and showed no interest in peer-revision in collaborative projects.

In terms of how peer-revision is performed, according to the survey, nearly half (45%) of the respondents would do a full check of the translation when revising others' work. 37% preferred to do a spot check by randomly selecting a part of the translation to revise. Both monolingual review and comparative revision were adopted in peer-revision. 5% read only the target text to check for any awkwardness or difficulties in understanding, while 37% compared the target text with the source text to check for mis-translations or other content-related problems. The rest of the respondents used both approaches – for instance, some tend to do a monolingual review first, and if there were too many problems, he/she would consult the source text and do a comparative revision (23%); some like to do a comparative revision first to check the content, and then a monolingual review to check the language problems.

By claiming to perform peer-revision in various manners, the survey result demonstrated a common awareness of potential content and language issues in crowdsourced translations. Given that participants are willing and

fairly competent in performing translation and revision tasks (6.4 discusses in detail the extent of their competence), implementing quality control methods in crowdsourcing translation appears to be a strong possibility.

4.6 Summary

This chapter painted a full-scale picture of *Yeeyan* participants' perceptions and experience regarding their own motivations for participation in crowdsourcing translation. First of all, my survey identified the two main demographic characteristics that the participants possess in order to engage in translation-related activities: higher education background, and L2 competence. The idea of 'natural translation' is thus supported by my data: although only 30% of the respondents had received formal translation training, the rest of them did perform translation-related activities using their L2 knowledge and accumulated experience in *Yeeyan*.

The most significant contribution of the questionnaire survey regards the motivations and de-motivations that may influence participants' behaviours in the crowdsourced environment: my data suggested that individuals are usually driven by multiple reasons both intrinsically and extrinsically, but extrinsic motivations are perceived as more important than the intrinsic motivations. The primary motivator is current and future rewards, particularly for the enhancement of human capital, despite altruism and enjoyment being the most powerful widely-acknowledged intrinsic motivators. The *Yeeyan* crowd's top two motivations – 'to improve language abilities' and 'to gain translation experience' – suggest that many participants treat the platform as a learning environment, which is worth considering by both crowdsourcers and academics. Moreover, this chapter identified de-motivations and other factors that may negatively influence participation, which include settings and properties of the crowdsourcing translation project, the communication and feedback mechanism, personal factors, conditions of the crowdsourced environment, and the censorship system in China.

More importantly, the analysis of the survey result revealed the relationship between motivations and behaviour: participants who are primarily motivated

by extrinsic reasons are likely to invest more time and effort into crowdsourcing translation than those motivated by intrinsic reasons – extrinsic motivations reinforce the participants' sense of responsibility and ethics.

Finally, the participation perspective of the survey showed that *Yeeyan* is a crowdsourcing initiative which depends on *integrative* contribution with *high* accessibility of peer contributions. Designing the collaborative platform from such a perspective matches the participants' motivations, as well, who usually play multiple roles to support the full functioning of the community. Quality control in *Yeeyan* replicates the professional TEP mechanism and other methods used in the professional industry, though in actual practice they appear to be more chaotic, with overlapping and repeated stages. The survey data regarding participation is self-declared, yet is still a useful dimension to consider when analysing the additional empirical data collected through observation, as we will see in the following chapters.

Chapter 5 Findings of the observation stage: 1 – workflow, productivity, and the learning perspectives

This chapter presents the primary findings of the observation stage of my project. It elaborates on the workflow of each project to give an overview of how translation projects progress in *Yeeyan*, and then discusses the productivity of participants in this particular environment. It also analyses the organisational structures of the crowdsourcing translation projects. Finally, the learning perspective of crowdsourcing translation is explored based on the workflow and data collected from the follow-up interviews in the *Gutenberg Project*. By emphasising the participation and organisation of crowdsourced translation projects from empirically-proven data, self-declared findings from the questionnaire survey presented in Chapter 4 are examined here, as well.

5.1 Data description – workflow mapping

Dunne (2011) argues that the most enduring concept model in project management is the triangle system, which indicates three fundamental factors: cost, time and quality. In the current market for translation and localisation services, time is arguably the most critical constraint. At the 2011 Localisation World conference in Barcelona, Derick Fajardo from *Nuance* observed that ‘cost rules, quality is assumed, but in the end, schedule wins’ (Beninatto in Dunne, 2011:120). ‘On-time delivery’ is considered as the most important metric for clients (ibid). Obviously, regarding the translation projects in this research, cost is not an issue that needs to be addressed. As indicated in the characteristics of observed projects in section 3.2.1 of the Methodology chapter, participants in the four *General Collaborative Projects* work voluntarily, for free; at the same time, remuneration in the *Gutenberg Project* is success-based – participants gain royalties based on the sale volume of the published translation.

As for the ‘time’ factor, all the observed projects went beyond the initially agreed deadline*: Health Project I* was 4 days late; *Health Project II* was 2

days late; *Business Project I* was 4 days late and *Business Project II* was 2 days late; the *Gutenberg Project* was 6 months behind schedule. Yet going over the allocated time does not seem as severe as in the professional industry: the text editor who called for the translation project did not show any dissatisfaction about the missed deadline, neither did she adopt any coping mechanisms to speed up the process when the first signs that the projects would take longer than initially planned appeared.

The discussion of 'time' is closely related to schedule, which is often associated with the design of the workflow. In the professional industry, a traditional TEP approach has been favoured for years. TEP refers to the process of translation (translation of the text), editing (bilingual revision of the translation) and proofreading (monolingual review of the target text). Language services providers and translation services providers may adopt various workflows, but they all need to guarantee that the implemented workflow and the resulting end-product are functionally and qualitatively equivalent to TEP. As mentioned in section 2.5.2 of the Literature Review chapter, the current widely-accepted quality standard – ISO 17100:2015 (International Organisation for Standardisation, 2015:10) specifies that the translation production process 'shall at least include the translator's overall self-revision of the target content for possible semantic, grammatical and spelling issues, and for omissions and other errors, as well as ensuring compliance with any relevant translation project specifications' and '[t]he TSP shall ensure the target language content is revised. The reviser, who shall be a person other than the translator […] shall examine the target language content against the source language content for any errors and other issues, and its suitability for purpose.' (ibid.)

During data collection, I observed that the practices of translation projects in *Yeeyan* actually conformed to the service requirements in ISO 17100 to a large extent, but also displayed individual characteristics. A similar TEP workflow was adopted. Revision as both self-revision and peer-revision was performed by participants spontaneously or through coordinated effort. Other quality control methods, such as using a shared glossary list, peer discussion for problem-solving, and a domain expert's proofreading also featured in the

observed process. In order to visualise the data, I used workflow mapping techniques for analysing the progress of the projects and for highlighting the individuals' performance, which is linked to the power relationships of the social network analysis (SNA) which will be mentioned in Chapter 6. On the one hand, this methodology is used to reflect how the translation tasks are collaboratively fulfilled in the crowdsourced environment; on the other hand, it is used to verify the result from the SNA of the interaction data.

To begin with, observation showed that the general workflow of a collaborative project in *Yeeyan* followed the same traditional TEP approach used in the Language Services Industry. In the setup stage, the project manager – PM (in the observed *General Collaborative Projects*, a recruited text editor in *Yeeyan* acted as a PM; in the *Gutenberg Project*, a PM was recruited by the *Yeeyan* editor in advance) – sent a recruitment announcement, inviting *Yeeyan* members to volunteer to participate. The PM then divided the source text into several parts with similar wordcounts according to the number of volunteer participants. Translators then signed up to a specific part and started translating. The PM usually started by giving a short project overview complete with instructions which indicated the workflow, deadline, how to use the text editor (i.e. the collaboration platform), and asked participants to create a shared glossary list while translating to ensure consistency. The next stage can be considered a classical translation stage. Then, the revision stage included self-revision and peer-revision, which were not specified as compulsory in the PM's instruction in *General Collaborative Projects,* whereas revision was clearly required in the *Gutenberg Project*. The finalisation stage was often performed by the PM, who dealt with text formatting aspects, and then published the translation on the website. However, if there was a domain expert in the group, a full text proofreading would be performed by this person before publishing.

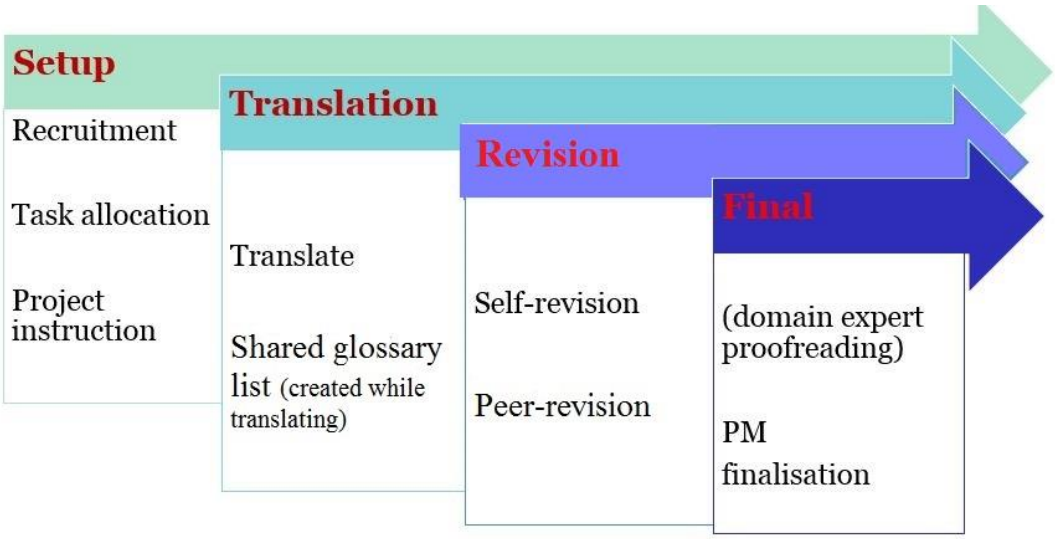

Figure 17. General workflow in a collaborative project in *Yeeyan*

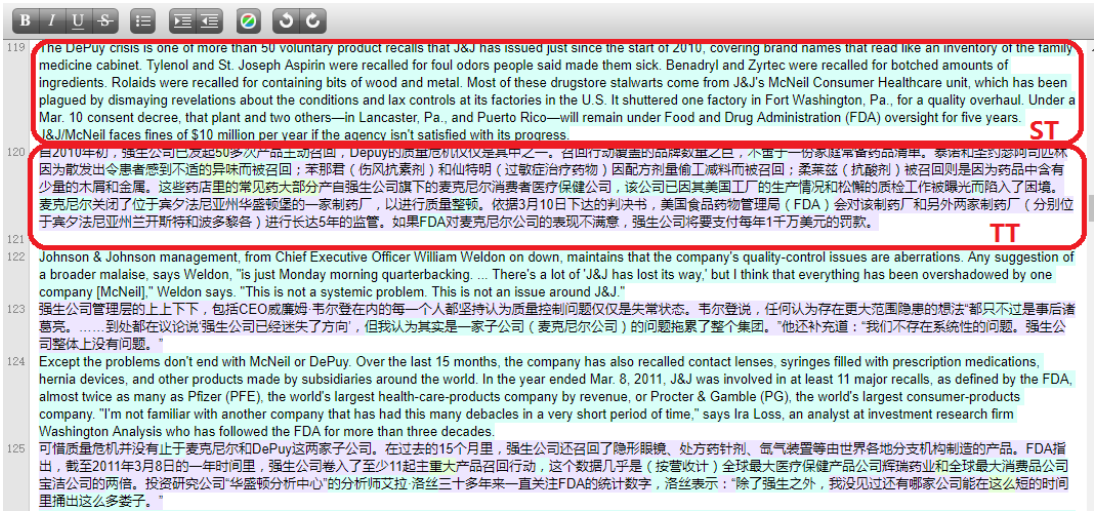

Figure 18. Screenshot of the edit pane for a general collaborative project (*Sync.In*)

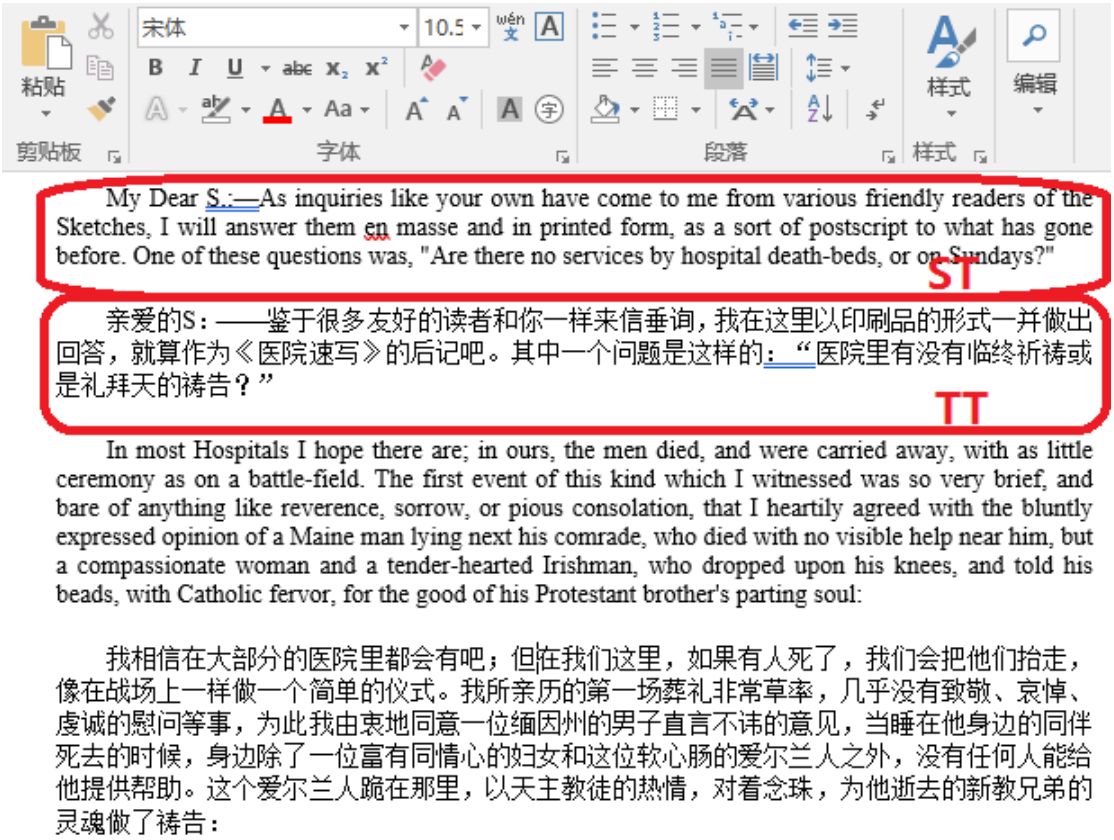

Figure 19. Screenshot of the edit pane for the *Gutenberg Project* (*Microsoft Word*).

Figure 17 maps the general workflow of a collaborative project in *Yeeyan*. Because the layout of the edit pane that carries all the editable tasks displays both source and target text segmented at paragraph level (see Figure 18 and Figure 19), there is no possibility to track whether the participants performed revision bilingually or monolingually. In the thesis, the term revision is used to refer to edits made after the translation is finished.

5.1.1 Workflow mapping of *General Collaborative Projects* – overlapping tasks and random working order

PMs in three projects specified that the translation should be revised before finalisation (in *Health Project I, Business Project II*, and *Gutenberg Project*). The remaining two projects did not have any requirements regarding revision: translation was the only task for participants in the initial project instructions. Although there was the pre-defined TEP workflow, the actual practice in the *General Collaborative Projects* did not stick to a sequential order. First of all,

the translation and revision stages were conducted simultaneously. Secondly, participants did not work in a top-down order – to be more specific, revision (both self-revision and peer-revision) was performed in advance. Some participants took the initiative to revise either their own parts of the text or others' assignments. There was no regular pattern in the revision stage.

The second finding that contradicts the conventional top-down order refers to the fact that participants started either translation or revision randomly. Normally, the logical order when translating a large text is to work from the beginning to the end for reasons of coherence. In the observed projects, the workflow did not follow the sequential order of translating part 1, then part 2, part 3, etc. until the final part. The same situation was also found in the case of revision.

The workflow of each observed project is mapped in Figures 20-24 for the purpose of finding out the characteristics of collaborative projects in the crowdsourced environment. The participants' identity is anonymised by using T for translator, R for reviser, and PM for project manager. The anonymisation code for the translators' identity corresponds to their task assignment – T1 is responsible for the translation of part 1, T2 for part 2, T3 for part 3, et cetera. One thing worth pointing out is that the Rs (revisers) are the supplementary participants, meaning the ones who joined the projects at the revision stage and did not engage in the translation task.

In Figures 20-23, for the four *General Collaborative Projects*, I described the workflow under the unit 'day' (D) along the *X* axis, which does not represent the actual productivity of an individual for a particular linguistic activity. The 'productivity' aspect is dealt with in section 5.2.1. Stages in the workflow are reflected along the *Y* axis. The pre-defined deadlines were reflected under the *X* axis. The most salient finding of the workflow mapping is that the PM mostly contributed during the setup and final stages, and sometimes during the end of the revision stage. The PM prepared the project by dividing the source text and specifying the requirements on the first day. It took around 1-

2 days to have all the participants sign up to their tasks.

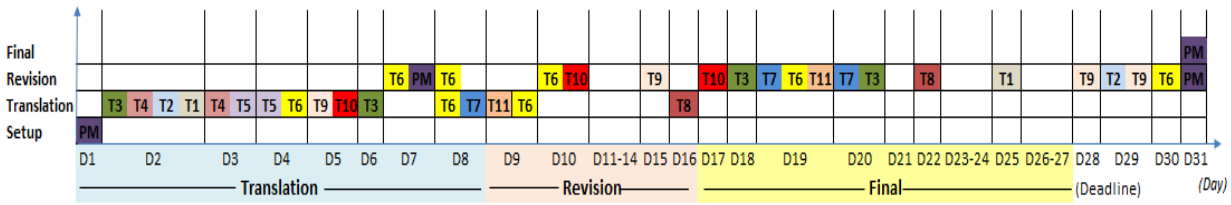

Figure 20. Workflow mapping of *Health Project I* (*HP I*)

In the *Health Project I* (Figure 20), translation started from the second day after the project was initiated. T3 was the first one that started translating and needed two days to finish the first draft (on Day 2 and Day 6). T1, T2 and T4 also started their translation on the second day, and T1 and T2 finished the translation task that same day. T4 continued the translation during the following day, which was when T5 first started the translation task. T9 and T10 started translating on Day 5. T3 continued the translation on Day 6.

On Day 7, T6 revised the translation for T1, T2 and T3; the PM also performed part of the revision for T1 and T10. The translation work continued on Day 8 and Day 9 (carried out by T6, T7, and T11). From Day 10 to the end of the project, revision was the most significant linguistic stage which occupied 2/3 of the project schedule. In total, T6 revised the highest number of words: for T1, T2, T3 and T4 on Day 7, Day 8, Day 10, Day 19 and Day 30. T10 revised T9 and T11's translation on Day 10 and Day 17. T3 checked T5's translation work on Day 18. T7 and T8 cross-checked each other's translation on Day 19, Day 20 and Day 22. T2 revised for T6 on Day 29. T9 revised for T10 on Day 15; and T11 on Days 28 and 29. The PM did a final revision for T1 on Day 31.

The workflow mapping demonstrates that not all participants performed the revision activity. T6 contributed the most during revision, which made her the *lead* participant. T6 was also the *key* actor in the social network analysis in Chapter 6 (section 6.2.3). According to the interaction between participants, T6 claimed to be the 'domain-expert' in the group and volunteered to do a final check of the translation, which explained the extensive effort she

invested.

As mentioned earlier, *Health Project I* missed the pre-defined deadline by 4 days. Figure 20 demonstrated that there were 9 days in total in which no activity was performed. In addition, three translators (T6, T8 and T11) missed their own deadlines for translation. As a result of these delays, the revision stage was also postponed. However, at a general level, the schedule for each stage of the project did not seem to be efficient, with the PM scheduling 11 days for finalisation – much too long for text formatting and publishing given that the translation was only published on the *Yeeyan* website for knowledge dissemination.

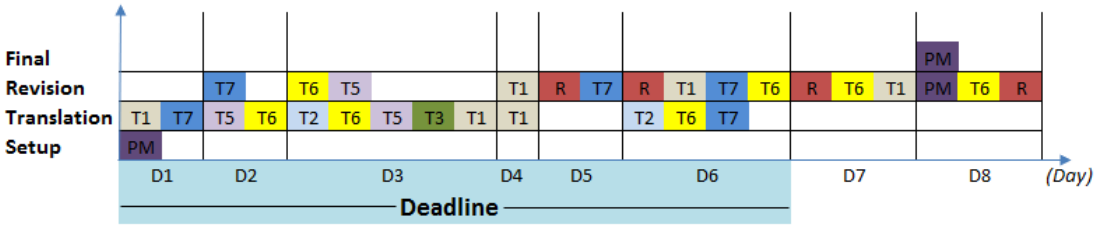

Figure 21. Workflow mapping of *Health Project II (HP II)*

Figure 21 maps out the workflow of *Health Project II*. In this project, T4 dropped out the project without any notice. The PM did not provide instructions on the workflow (i.e. whether revision was needed) and merely stated a deadline when the translation needed to be finished. On Day 6, T2, T6 and T7 did the translation initially allocated to T4. Part 4 is used to represent the text that was initially assigned to T4. Therefore, there is no sign of T4's engagement in the workflow mapping.

T5, T6 and T7 were the ones self-revising on the next day after their translation was finished (Day 2 and Day 3). T1, T6, T7 and a third-party reviser (R) actively performed the revision work. T1 revised T2's translation on Day 4, Day 6 and Day 7. On Day 5, T7 checked T6's work, while R revised for T7. T6 checked translation produced by T1 (on Day 6), T2, T3, and T5 (on Day 7), and Part 4 (on Day 8).

The extent of the participants' involvement in the revision activity reflects the contribution of participants in the project, which also corresponds to the result

of the social network analysis detailed in in Chapter 6 (section 6.2.3). T6 is the *lead* participant who contributed the most effort in revision in *HP II*, in the meantime, T6 is also identified as the *key* actor in the group discussion by acting as an information provider. Because T2's translation appeared to be the most problematic one in terms of quality (see section 7.1, in Chapter 7), which was a widely agreed fact among the participants (indicated in the group discussion), T2's work was repeatedly revised.

As can be seen from Figure 21, *Health Project II* was delivered two days after the PM's required deadline. However, it was the revision effort on the last two days that ensured the full text was of adequate quality. Although the PM did not specify revision to be included in the project, participants spontaneously performed this activity.

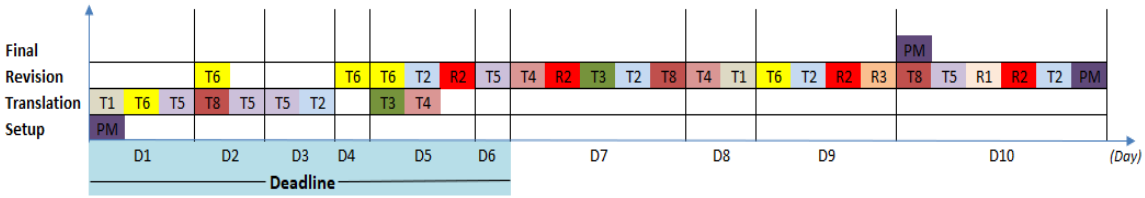

Figure 22. Workflow mapping of *Business Project I* (*BP I*)

The workflow of *Business Project I* is mapped in Figure 22. Similar to *Health Project II*, the PM only specified a deadline for the project which eventually was four days behind schedule because of major revision work at the end. T1, T5 and T6 started translation on Day 1. T6 performed self-revision on the next day after her translation was finished (Day 2). T5 translated continuously during the first three days of the project, because T5 translated part 5 and part 7 on his own. T3 and T4 did their translation a bit later than the others (Day 5), which was the initial planned deadline. T5 and T3 also conducted self-revision (on Day 6 and Day 7, respectively).

All translators in this project performed the revision activity. Starting from Day 5, T6 revised T1 and T4's work, T2 for T3, and R2 for T2. On Day 7, T4 also revised for T1, R2 continued the revision for T2, T2 then checked for T4's work, and T8 checked for T6. T4 revised for T3, and T1 revised for T8 on

Day 8. Next, several contributors performed an extensive amount of revision: T6 revised for the works of T1-3 and part 7; T2 checked for T6, R3 for part 5, and R2 for part 7 on Day 9. On the final day, T5 revised the translation produced by T4 and T8, R2 for T1 and T3, T8 revised part 7, T2 revised for T8, R1 for T2, R3 checked part 7, as well, and the PM revised for T1. Revision activities in this project were mostly repeated. Each part had at least two participants' revisions. A third-party reviser – R2, alongside T6 and T2, stood out to be the ones who contributed the most during the revision task: they were the *lead* participants in terms of contribution quantity. T2 is recognised as one of the *key* actors in the interaction network in *BP I* (see section 6.2.3 in Chapter 6).

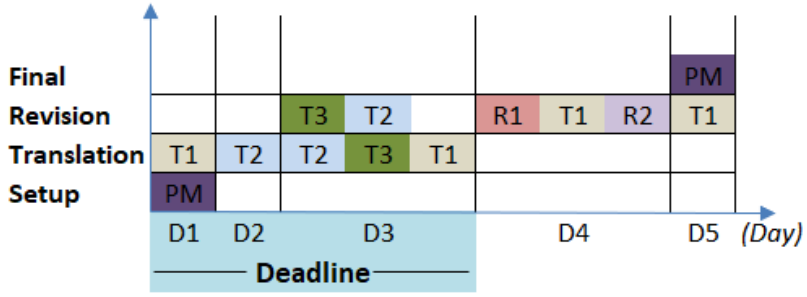

Figure 23. Workflow mapping of *Business Project II* (*BP II*)

Business project II is a relatively small project compared with the previous ones; Figure 23 maps its workflow. Because it was initiated for beginners to practice translation skills, the wordcount of the selected source text is only 962, divided into two parts: T1 and T3 translated part 1, while T2 translated part 2. The PM specified that the translators should revise each other's work, and for training purposes, two senior members in the community were invited to feedback through their revision, as well. Only a final deadline was set. The project missed the deadline by two days. The translation task was finished right on time, but the revision was performed mainly on the two extra days. To be more specific, T3 revised T1's translation, while T2 revised the entire Part 1 on Day 3. A third-party reviser R1 checked part 1, T1 revised part 2, and R2 revised the whole text on Day 4. As shown in Figure 23 , linguistic activities performed by both translators and the senior reviser were quite

evenly distributed along the timeline, except that T1 did a final check of the full text on the last day. All translators who had signed up for the task revised each other's work according to the PM's instructions. Moreover, the social network analysis in section 6.2.3 (Chapter 6) shows that T1 was the most active participant (i.e. key actor) in the interaction board, followed by T3, and then T2.

In general, the workflow mapping demonstrates that in open, voluntary crowdsourcing projects, participants spontaneously sign up for individual tasks, perform them fully in the overwhelming majority of cases, and some participants even do extra work revising others' translations. In the absence of much stricter processes such as the ones in the traditional TEP approach, participants choose to perform the translation and revision tasks in a somewhat more random order. Although for some critics crowdsourcing equals a total lack of planning and regulation, participants in the observed projects showed the awareness that a collective project is not finished until each individual part is both translated and revised. Hence, I define such phenomenon as **user-driven flexible workflow** as it largely depends on the participants' initiative alongside a large degree of spontaneity in performing translation-related activities.

Very interestingly, the workflow mapping contradicts in one of the self-declared participation patterns demonstrated in the questionnaire survey: while 95% of the respondents declared to self-revise their work, only a few participants in the observed projects performed self-revision.

On the other hand, peer-revision was the most widely selected quality control method in the questionnaire, and in the observed projects each participant's translation was revised by other participants (with occasionally some parts being repeatedly revised). What appears rather difficult to predict and control, and at the same time what is very different from the professional Language Services Industry is the amount of revision which every translated section receives. Very often one part has revision edits by multiple people, while another part is only revised by one person. Sometimes, one part of the project was left unchecked, which would then prompt a translator to ask a

teammate in the chat board to revise the work (a process discussed in Chapter 6).

5.2 Data analysis

5.2.1 Productivity assessment: translation VS. revision

The previous section (5.1.1) has mapped the workflow of collaborative projects in the crowdsourced environment. According to the data gathered in my experiment, revision takes longer than translation. On the one hand, since not all participants started working on their task right after the project was initiated, this caused a delay in the progress of the project. On the other hand, translation and revision were not performed in an organised manner in that some participants would translate/revise on one day, while others would work on another day. For example, there were 9 days in *Health Project I* (*HP I*) in which no activity was performed. Therefore, the schedule cannot represent the productivity of the 'crowd'.

Unexpected developments can also influence the workflow – for example, in *Health Project II* (*HP II*), one participant (T4) dropped out for no apparent reason and left his/her part untranslated. Others (T2, T6 and T7) had to take over the task four days after the start of the project.

Moreover, it is necessary to note that participants work voluntarily in their spare time, and that the source text has certain difficulties for amateurs. Judging from the interaction record and edit history, most of the 34 translators work during late afternoons and nights; at least two of them were students, and 2/3 of them were working full time.

An interesting research question therefore can be: what is the productivity of the 'crowd' in such a less-governed collaborative project?

Individual participants' edits on the collaboration platform were retrieved to assess their productivity – how much time had each participant spent on the two linguistic activities: translation and revision? I recorded and analysed the time (minutes) people spent on translation related to the unit 'word count' in each allocated part of the source text, and also the time they spent on

revising another's translation (minutes and how many words were revised).

Table 12 presents the mean and standard deviation for each activity in the four *General Collaborative Projects*. Because participants in the *Gutenberg Project* (*GP*) chose to work privately using *Microsoft Word* instead of the collaboration platform, the time each participant spent on related linguistic activities is not trackable. Consequently, the productivity assessment only looks at the four *General Collaborative Projects*, as they represent the most original form of crowdsourcing with properties of open access, free volunteer participation, and minor supervision.

Table 12. Productivity assessment across *General Collaborative Projects*

Project Name	**Translation** Mean (SD) (minutes)	Word Count Mean (SD)	Number of variables	**Revision** Mean (SD) (minutes)	Word Count Mean (SD)	Number of variables
HP I	151 (63)	1099 (232)	n=7	95 (50)	877 (323)	n=10
HP II	106 (32)	733 (191)	n=7	83 (38)	592 (218)	n=10
BP I	118 (27)	756 (102)	n=8	84 (63)	713 (138)	n=22
BP II	115 (15)	473 (25)	n=3	98 (42)	688 (217)	n=5

Four participants in *HP I* and one in *HP II* did their task outside the platform and copied their translation to the edit pane, which makes 5 null variables. Table 12 indicates the performance of trackable and valid variables. The number of variables for revision refers to the rounds of revision done by individual participants for each part.

In the professional industry, a widely accepted productivity standard for a translator is 2,000 words per day (8 hours) (Language Realm, N.D.), and approximately 1,300-1,500 words per hour for a reviser (STP, N.D.). For the purpose of comparison, I also included the average productivity of trainees collected from *Students' Team Project*[20] from University of Leeds – Centre for Translation Studies *MA programme in Applied Translation Studies*. When evaluating productivity, several influential factors need to be considered: the difficulty of the source content, the quality of the source content, the available

[20] For details of the *Students' Team Project,* please see section 6.3.1.

resources, the quality of these resources (TAUS, 2015), and the quality of the translation when assessing revision productivity.

In the case of these four *General Collaborative Projects*, first of all, the source text analysis in section 3.2.4.1 Methodology Chapter, estimated that *HP I, HP II*, and *BP I* are challenging for amateurs in the crowdsourced environment. Secondly, the widely-accepted level of translation productivity refers to producing 'professional'-standard quality. But in the context of crowdsourcing translation, these texts are for knowledge-sharing, which makes them informative texts. According to the functional theory, the target informative text should 'focus on transmitting the factual content and terminology and not worry about stylistic niceties' (Reiss in Munday, 2012:114). With all these considerations in mind, the average productivity for translation is actually comparable with the professional industry (see Table 13). In terms of translation, the **average productivity** of the 'crowd' may be even faster than trainees and professionals (around 2.3 hours, 2.5 hours, 2.7 hours, and 4 hours per thousand words in *HP I*, *HP II*, *BP I*, and *BP II*, respectively). One must note that *BP II* was set up specifically for beginners in the *Yeeyan* community, which explains a relatively lower translation productivity. Also, translations produced in such short time by amateurs can be of **poor quality** that needs vigorous revision. Moreover, one must also take into account the possibility that individual translators may have spent time outside the platform to do research work or to think how to translate one segment, which is not shown on the platform.

Table 13. Average productivity across projects with reference to *Students' Team Project* and the *Professional Industry*

	Average Productivity (approximate) (hours per thousand words)					
	HP I	*HP II*	*BP I*	*BP II*	*Students' Team Project*	*Professional Industry*
Translation	2.3	2.5	2.7	4	4	4
Revision	1.8	2.3	2	2.3	1.5	0.7

More importantly, the data demonstrates that **the crowd's productivity of revision is much lower than** the revision productivity of trainee translators, as well as the expected standard in the professional industry. One reason might be that untrained non-professional participants are able to produce draft/poor level quality quickly, but it takes much longer and more effort to revise the translation, as suggested by the workflow mapping. This matches the current situation of crowdsourcing applications on the market: easy to get the needed raw service, but difficult to improve it to make a quality, useful product (Rajala et al., 2013). When annotating the participants' performance, I found that they do not work continuously as professionals do on a daily basis, mainly because this is spare-time activity performed by people according to their free will. For example, some work for approximately 20 minutes on the platform, and then come back to work for a short period several hours or days later.

Another finding of the experiment is that that the edits on the platform for the translation task are much more intensive than for the revision. Participants tend to focus more on the translation by putting more effort on a daily basis, meaning that the translation is a centralised activity, while revision receives less focus. It is another uncertain factor that influences the productivity assessment in Table 13.

In general, the result of the productivity assessment indicates that the productivity of both translation and revision in the crowdsourced environment is sufficient for a knowledge-sharing task – **producing a revised translation on a volunteer and less-governed basis under the condition that the deadline is not the priority of the project**.

5.3 Discussion

5.3.1 Comparison between the user-driven flexible approach and the waterfall approach

‘Autonomous’

From the workflow mapping in the four *General Collaborative Projects*, it can be argued that, in the crowdsourced environment, there is no clearly-defined hierarchical organisational structure as in the professional industry. ‘Autonomous’ is an often-emphasised feature of crowdsourcing. It refers to the premise that participants are able to perform the task according to their free will. Overlapping tasks and spontaneous multiple-role changing which were demonstrated through the workflow mappings prove the existence of the feature ‘autonomous’ in the translation projects done for knowledge sharing. There were areas – such as explicit allocation of tasks – in which PMs did not intervene excessively, leaving it to the group to choose tasks independently. This ensured a democratic, open environment for the participants to contribute to without the pressures that may influence or de-motivate one’s participation in volunteer work. It is the kind of environment that people need in a highly competitive society. On the downside, less monitoring can lead to a situation of repetitive work, which was observed in all projects: some translations were revised several times, whereas others were initially left almost unedited (the translator had to appeal for revision in the chat board).

A crucial finding of my research is that, while **repetitive work** may be considered as a waste of labour in the professional industry, **in the crowdsourced environment it is generally a necessary undertaking**, because less-experienced participants may introduce more errors when revising instead of correcting them. From the statistics obtained after the error annotation element of the experiment, most of the errors were being corrected through several rounds of revision (discussed in 6.4). For the *Gutenberg Project*, a different procedure was used, which makes the notion of ‘autonomous’ less visible in this specific kind of project. Because the

project is set for publication that requires professional-level quality, some restrictions were applied, as detailed below.

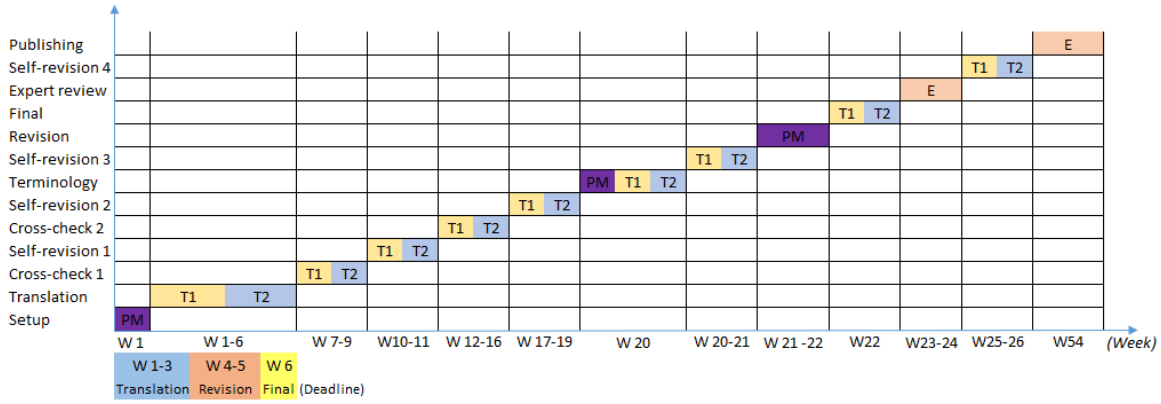

Figure 24. Workflow mapping of the *Gutenberg Project* (*GP*)

Figure 24 maps the workflow of the *Gutenberg Project*. Under the time unit 'week' (W), if a project aims for publishing quality, more procedures are implemented (e.g. self-revision, cross-check, PM full text revision and expert review). The term 'cross-check' is used to make the distinction between the revisions made by the translators and the PM. The *Gutenberg Project* followed a strict **waterfall workflow**, in which every stage has its own goals and, after one stage was completed, the process moved to another. During the observation, I found that the translation of each part was performed in a sequential order, and revision started after the completion of the entire translation.

According to the follow-up interview with the PM, *cross-check 1* aimed at improving the accuracy of translation and solving issues that the translators were not confident with. Self-revision(s) follow(s) cross-check(s), and in that stage the translators had to decide whether the suggestions/corrections made by team members were acceptable and to make further revisions as after the time gap the translator may have different thoughts. *Cross-check 2* was meant to deal with the language issues to make the translation sound more fluent and natural according to the target language conventions. *Self-revision 3* was to ensure the terminology consistency in accordance with the shared glossary list built by the PM (the terms were agreed by all participants

during the group discussion). Prior to the finalisation, the PM performed a full text revision which significantly improved the translation quality (see section 7.2, Chapter 7). Different from the four *General Collaborative Projects* and the professional industry, a finalisation was performed by the two translators in line with the project requirements and guidelines for publication specifically for the *Gutenberg Project*. Afterwards, the finalised texts were submitted by the translators for expert review. Because this was conducted privately, I did not have access to the files (error annotation only involved the versions before the expert review, which I compared against the published version which I used as the 'gold standard' for reference). The PM stated in the follow-up interview that the suggestions provided by the expert reviewer mainly addressed style issues, and the translators self-revised the full text accordingly. In general, this result corresponds to the findings in the *General Collaborative Projects* – revision takes much longer than translation, which makes the project exceed the initial deadline.

As a solicited, strictly-governed project, it applied the traditional **waterfall approach** when being scheduled. The *Gutenberg Project* progressed step by step abiding by the traditional TEP approach. Prior to the setup of the project, there was an open recruitment announcement that called for participation. A sample test was given with around 300 words of text for each participant to translate. According to the interview with the PM, she indicated that two core principles governed the selection of suitable participants: 1. good translation quality; and 2. people with a similar translation style. All this preparation work seems to imply that it is less autonomous to participate in the *Gutenberg Project*. Out of 44 applicants, two were chosen based on the result of the sample test, so one may argue that participants do not actually have the freedom to decide whether and how they wish to contribute, particularly in the preparation stage of the project because of the professional-equivalent quality requirement.

When it comes to the translation and revision activities per se, participants have full freedom in deciding how to translate, whether to accept a revision or not, and when to submit, which are reflected in the communication between translators and the PM. There is a large negotiation space in terms of the

decisions that need to be made regarding the substance of the translation. Translators can oppose each other's suggestions and discuss to solve disagreement, which is reflected in the large amount of information-related dialogue acts I observed (see Figure 26, Chapter 6). According to ISO 17100:2015, whether the changes made during the revision are implemented is decided by either the reviser or the translator depending on who does the finalisation (STP, N.D). Usually, there is no negotiation between the two parties. From this perspective, it seems that participants in the *Gutenberg Project* have more power in deciding the implementation of changes made to their translation. However, as shown in Figure 24, an expert review is included towards the end of the project. It is done by the professional editor in *Yeeyan* who decides whether the translation meets the quality standard for publication. Translators had to re-revise according to the feedback provided by the professional editor. Therefore, I argue that volunteers in a strictly-governed project, e.g. *Gutenberg Project*, have less autonomy than participants in a less-governed project, e.g. *General Collaborative Project*.

'Organisational structure' – Network structure and Hierarchical structure

Table 14. Labour distribution across the observed projects

		HP I	***HP II***	***BP I***	***BP II***	***GP***
Client		PM	PM	PM	PM	End-client (Publishing House)
Organiser	PM	1	1	1	1	1
Primary Participants	Translators	11	7	8	3	2
Supplementary Participants (Gatekeeper(s))	Revisers	None	1	3	2	None
	Professional editor	None	None	None	None	Yes

The labour distribution in Table 14 shows an organisational structure of collaborative translation projects in the crowdsourced environment. It can be called '**network organisation**', which by definition means that autonomous actors collaborate to supply a product or service (Aalst and Hee, 2004), or a **virtual organisation**. An extended feature of 'autonomy' is identified,

meaning that there exists no formal permanent (employment) relationship, which enables an individual to choose whether or not to carry out a particular task. The actors required to perform each task therefore must be recruited individually on each occasion (ibid). This highlights the autonomous feature from a new perspective: *personnel*, meaning a labour force with more mobility. In the crowdsourced environment, therefore, autonomy is reflected from two perspectives: one refers to the sphere in which the participants can decide their own activity; the other refers to a mobile personnel relationship between the organiser and participants.

In the example of *Yeeyan*, the organiser calls for translators every time when a new project is set up. For these projects aiming at knowledge-sharing, there is no participation entrance requirement – this merely depends upon the *Yeeyan* members' personal will. In addition, 'autonomous' is further featured through the allocation and acceptance of work. Normally, in the professional industry, under the top-down working condition, most people's work is assigned or outsourced to them by their principals, either by the boss or direct client. There is usually a communication protocol between the principal and the actor, such as task specifications, quote, assignment, confirmation, purchase order, etc. A crowdsourced project skips these formal communication procedures. As described in the workflow mappings, the PM did not assign tasks to specific participants; instead, the PM left the choice to the actors – allowing participants to sign up to the desired task involving a specific part of the source text, which is an encouraged spontaneous process for the allocation and acceptance of tasks.

Compared with the **user-driven flexible** workflow in the *Yeeyan General Collaborative Projects*, the **waterfall workflow** advances strictly under the PM's supervision. This represents a typical **hierarchical organisational structure** which is also known as a 'tree' structure (Aalst and Hee, 2004). The 'leaves' of the tree usually represent the groups of staff. The 'branches' demonstrate the authority relationships: the person at the start of the branch is authorised to make demands and to order work from the staff at the end of it. The higher 'leaves' can influence the structure below them by allocating staff with different functions (ibid). In the *Gutenberg Project*, the PM decided

the procedures, task allocation and deadline of each procedure according to the progress. More importantly, primary participants signed a contract with *Yeeyan* which set a commitment with the main labour force. Primary participants shall abide by the contract which specifies the rules and remuneration. They shall work as a collective project, meaning that primary participants are mandated to perform both linguistic activities – translation and revision – until the output is approved for publishing. Therefore, in a solicited, strictly-governed crowdsourcing project, under the waterfall workflow, they volunteered to take part in the project, but more restrictions are applied when performing activities.

'Span of control'

Another important aspect associated with the organisational structure in project-based environments is the span of control: who takes a leading role in organising and managing the project?

Looking at the staff allocation in the two categories of observed projects (Table 14), in the four *General Collaborative Projects* that aim at knowledge-sharing, actors include one PM, primary participants who signed in as translators, and supplementary participants who worked as third-party revisers (their presence may vary depending on the project). On the other hand, in the *Gutenberg Project* that aims at publishing, the actors are the end-client (publishing house), the PM, primary participants who signed in as translators, and supplementary participants who are the salaried expert proofreaders.

In the workflow mapping, the former four *General Collaborative Projects* were processed in a user-driven flexible approach that firmly depended on the spontaneity of the primary participants. They self-organised the workflow and spontaneously performed linguistic activities. The same applies to the volunteer third-party revisers who joined the project without any formal/informal invitation from the organiser or the primary participants (with the exception of the *Business Project II)*. Moreover, revisers self-selected which part they wished to revise, while sometimes their action was called for through others' revision request (RR) in the chat board (discussed in Chapter

6).

In a network structured crowdsourcing project, after the organiser initiates the project, primary participants take a leading role in terms of control power due to the ‘autonomy’ feature of crowdsourcing, followed by the supplementary participants (volunteer third-party revisers) because their engagement is more flexible than the primary participants. By contrast, in a hierarchically structured crowdsourcing project like the *Gutenberg Project*, the end client has the largest influence in the span of control in deciding whether the output is suitable for publication. A further professional editor can act as the gatekeeper to ensure quality by deciding whether the output needs revision, what kinds of revision are needed, and whether the revised translation can be considered a completed project. The PM shares some of the linguistic control with the professional editor, but the PM has more power from the personnel perspective. The PM decides who to recruit as primary participants and if third-party assistance is needed. Finally, the primary participants take control during the production process mainly for linguistic activities. In short, in the hierarchical organisational structure, more stakeholders are involved, and they focus on different perspectives in the span of control.

‘Project management’

Table 15 presents the tasks completed by the PMs in the two kinds of crowdsourcing projects in *Yeeyan,* with reference to the requirements of project management in ISO 17100:2015. Three tasks marked in italics and bearing an asterisk (*) are exclusive to the crowdsourced environment I researched. For ease of reading, the performance of the PM in each task is evaluated in the form of tick (✓) on a scale of two. Two ticks represent that the PM performed the same level of activity as required in the professional industry, while one tick represents that only part of the required activity was performed. The cross mark (✗) indicates that no activity was conducted.

Table 15. Attributes of project manager across the observed projects with reference to professional industry

Project Manager (PM)	*General Collaborative Projects*	*Gutenberg Project*	*Professional Industry*
Identifying project requirements	✓	✓✓	✓✓
*Call for contributors**	✓✓	✗	✗
*Preselection of contributors**	✗	✓	It depends.
Preparation	✓	✓✓	✓✓
Task allocation	✗	✓✓	✓✓
Delivering project requirements	✓	✓✓	✓✓
Monitoring schedule and deadlines	✗	✓✓	✓✓
*Revision**	✓	✓✓	✗
Communicating any changes of the project specification	✓	✓✓	✓✓
Answering translation and other queries	✓	✓✓	✓✓
Managing feedback	✗	✓✓	✓✓
Verifying the translation with compliance of project specifications	✗	✓✓	✓✓
Delivering	✓✓	✗	✓✓

This mapping reveals that the PM in the *Gutenberg Project* actually performed the required tasks at almost the same level as in the professional industry. One exception occurred in one of the crowdsourcing-exclusive features – the call for contributors – because the *Gutenberg Project* was formally initiated by *Yeeyan* in collaboration with the publishing house, the 'call for contributors' was announced by *Yeeyan* in an official format rather than by the PM in the usual informal forum post (in the *General Collaborative Projects*). Both the PM and translators were recruited through the call. However, in cases where the PM is selected in advance, he/she then is in charge of the selection of the translators. Another difference between the *Gutenberg Project* and a professional translation project lies in the PM's engagement in revision. Normally, in the Language Services Industry, the

PM's responsibilities do not include revision. My study found that the PM's contribution to the revision stage in the observed *Gutenberg Project* played a significant role in improving the translation quality (see section 7.2.4 in 6.4).

It is also worth pointing out that the PM in the *Gutenberg Project* made great efforts to prepare the project in its setup stage, including drawing a detailed schedule, writing task instructions for each stage of the project, and providing reference resources (i.e. terminology databases for foreign names and places in an *Excel* format, a translator's handbook, rules for adding notes to the translation, explicit publishing standards of writing numerals and punctuation, as well as a formatting guide). The PM also checked the status of each translator's task regularly to monitor and ensure compliance with the schedule throughout the project. When it came to the expert review, the PM also made sure that the feedback was delivered to the translators and that the translators were implementing the suggestions accordingly when revising.

As mentioned earlier, the project missed its initial deadline perhaps due to the lack of experience of the PM, who overestimated the translators' productivity and competence. However, the PM managed to motivate translators to finish the project even when repeated effort was required. According to the follow-up interview with the PM, she indicated that her primary motivation in participating in the *Gutenberg Project* was the opportunity to have her work published: '*it feels good to see my name on the cover of the translated book*'. Hence, the extrinsic motivation of *self-marketing* sustained her devoted time and effort in the project, which matches the findings regarding motivations and their influence on one's behaviour discussed in Chapter 4.

In the case of the *General Collaborative Projects*, the PM's tasks were simpler and more flexible, and also differed according to personal will. The PMs did not make specific requirements of the projects except the deadline and a shared glossary list. In two projects (*Health Project I* and *Business Project II*), the PM indicated that revision was needed; nevertheless, the extent of the revision work that the PMs were involved in was not comparable

to that in the *Gutenberg Project*.

All in all, the responsibility and influence carried by the PM in the *Gutenberg Project* are much greater than in the *General Collaborative Projects* in terms of quality assurance, but also general organisation which means that, in such projects, the selection of the PM needs to be done carefully, to ensure that she/he has both project management skills and good linguistic ability, as well as strong motivation.

5.4 A learning perspective

The crowdsourcing TEP-like workflow also includes a learning component, particularly for collaborative projects, given that the revision edits and comments are openly accessible to all participants. As discussed in section 2.2.5 in Literature Review chapter, 'deliberate practices' and 'feedback' are the core components that make crowdsourcing translation projects a learning environment. Self-assessment is a commonly used pedagogical method which encourages participants to evaluate their own work and the learning process, and identify learning points, as well as the strengths and weaknesses of the activities performed in order to make improvements (Andrade and Valtcheva, 2009). In this research, a mock self-assessment strategy was used to investigate whether collaborative projects in the crowdsourced environment facilitate learning.

After the finalisation of the *Gutenberg Project,* while the book waits for publication, semi-structured interviews were conducted to further explore the translators' and PM's experience and perceptions on participating in the project. During the interviews, I asked 'what have you gained from this project? Can you think of any improvements that you realised after the project?' Although no criteria were given to assess the learning process, the participants identified several weaker aspects that they have improved.

To be more specific, the PM in the *Gutenberg Project* (*GP*) indicated that her practice in '*selecting translators, preparing and monitoring the project, discussing and answering translation problems, managing translation quality, as well as communicating with all stakeholders*', **diversified her experience**

from many aspects. She **developed more skills** rather than merely being a translator.

Table 16 presents the answers provided by the two translators in the *GP*. In addition to the interview data, it also includes part of the introductory remarks written by the two translators which were subsequently published with the electronic copy of the book (T1's remarks were published as the *Translator's Afterword*, while T2's remarks were published as the *Preface*).

Both translators ascertained their improvement in language ability. During the interview, T1 declared that he had joined *Yeeyan* with the motivations of practising translation and personal interest. He spoke highly of the strict revision procedure in *GP* by acknowledging the suggestions provided by his team members and the external expert reviewer. He indicated, '*T2 and the PM helped **improve the accuracy** of my translation. We discussed the controversial issues in the revision stage. […] For **language problems** mainly, the discussion brought our **translation style** closer. It is good because, compared with the first draft, there are **lots of improvements** in the final version*.' Further reflecting on his own development, T1 emphasised his increased knowledge of English, and also acknowledged the benefit which the slower manner of reading for translation brings to the ultimate understanding of the source text.

Table 16. Self-assessment of learning reflected in interviews and supplementary material (the *Gutenberg Project*)

Interview T1	Translation
'嗯，至少英文有提高点，另外呢，也算更认识自己一些写作和行文上的不足，然后可以多读些现代作家或翻译家的文章来弥补一下吧。另外也可以认识一些翻译达人，来做好朋友。'	*'Yes, at least **my English has improved**. Besides, I **learnt my deficiencies** in writing and rendering [in translation]. I will **remedy** my limitations by reading more articles written by modern writers or those translated by professional translators. Also, I **got to know** some "expert" translators and made friends with them.'*
Translator's Afterword by T1	Translation
'身为读者，读了几次，有时觉得	*'I have read the book several times merely as a reader. Sometimes, I felt it*

写得极好，语言生动活泼，可谓满纸机锋俏语，让人流连再三，不忍释卷。可有时再读又觉得言语啰嗦，结构松散。文本并没有变化，只是心绪却很无常，而我们自己又常常做不了自己情绪的主人。但翻译却常常能让人静下心绪，慢慢鉴赏文本，发现些乍见之下无法看到的"真相"。'	*was really a good piece of writing. The vivid illustrations made it so interesting that I was totally drawn in and could not endure to close the book. Sometimes, I felt the language was a bit verbose and the structure was too loose. There is no actual change within the text itself, it is just my emotion that keeps changing. However, we cannot be the master of our own emotion. Translation, on the contrary, calms me down, gradually, I found myself **evaluating and appreciating the text**, and I began to **realise the "truth"** that I did not notice at first sight.'*
Interview T2	Translation
'收获非常多。对我来说越是让我头疼的文本事后对我的提高越大，另外就是合作，一是我从 T1，PM 那里学了不少东西，包括英文和中文，二是感受到合作的好处，例如 T1 的一段精彩译文我恐怕很久都想不出来。'	*'I **gained a lot** [through participation]. For me, the more troublesome the text is, the more **improvements** I can draw. Another benefit is from collaboration. On the one hand, I **learnt a lot** from T1 and the PM, including **both English and Chinese skills**; on the other hand, I realised **the benefits of collaboration**. For example, a paragraph of translation done by T1 was so fascinating that I could not come up with it even after days of consideration.'*
Translator's Preface by T2	Translation
'但开始着手翻译的时候又让人有点头疼，因为本书写成于美国南北战争期间（1861—1865），书中引用了当时流行的典故和事件，对于这段历史不是很了解的中国人来说，有时候真是一头雾水。另外，当时的英语作家喜欢用非常长的长句，有时甚至一个段落就是一句话，其中倒装、省略又比比皆是，很多地方都成了我们翻译中的拦路虎，为了最大程度还原原著的内容和文采，我们不得不通过查找资料、组织讨论、求助专家等方法解决这一个个难题。'	*'It was a bit of **headache** when I **first** started the translation because the book was written during the American Civil War (1861-1865). It quoted some stories and events at the time, which can be confusing for Chinese readers who do not know that period of history well. Plus, English writers at the time liked to use very long and complex sentences. Sometimes, a paragraph was made up of a long sentence, in which inversion and ellipsis occurred frequently. They became the main **obstacles** in translation. In order to restore the content and style of the source text to its largest extent, we made great efforts **to solve the difficulties** in **language rendering**, including research work, group*

	discussion and asking for help from linguistic experts, etc.'

T2 took part in the project purely because of personal interest in foreign literature. As indicated in the interview, T2 enjoyed reading foreign literature but he soon realised some of the translated versions were of poor quality. He then turned to read the original versions and finally decided to try to translate the work. Prior to the *Gutenberg Project*, T2 had completed two literary translation projects in another crowdsourcing initiative (both were formally published: one independent translation and one collaborative translation).

When I asked 'if there is any significant difference after each round of revision' in *GP*, T2 acknowledged that '*there is **visible improvement** in quality, for both of us. T1's **suggestions** are very **useful**. He helped me find out the problematic part, some of them are **content errors** due to my misunderstanding, a majority of them are in the **language** part, the way of rendering**. My problem** is that I cannot use Chinese properly after translating for a long time; my translation can sound very odd*.' Combined with the error annotation detailed in 6.4, the comparison between the sequential versions (cross-checked version and self-revised version) demonstrates that peer-revision helped spot errors and improve both accuracy and fluency, most of the corrections made by T1 having been accepted by T2. As illustrated in the *Preface* of the published book, T2 understood where his weaknesses were and what kinds of strategies can be used to solve these problems, which is direct evidence of the learning process in collaborative project.

The process of learning is iterative as participants' output continuously evolves, and knowledge is continuously shared and updated among participants (Tran et al., 2016). In this regard, the TEP translation workflow continuously reconfigures the language knowledge of the participants, especially when several rounds of revisions were implemented. The result of the self-assessment in the *Gutenberg Project* matches the findings of the questionnaire survey in Chapter 4: both invisible and visible rewards were gained through participation (i.e. the improvement of translation and language ability, as well as enhanced reputation and peer recognition). Consequently, I propose that **an iterative learning process be embedded**

in crowdsourcing translation projects, with the aim of facilitating user growth in both personal competence and additional future developments.

Revision edits provide elaborate individual feedback to participants, which is not always feasible to implement in a classroom setting because few instructors have the luxury of regularly responding to each student's work (Andrade and Valtcheva, 2009). However, the integrative contribution setting and high accessibility of peer contributions embedded in the TEP-like workflow in crowdsourced projects guarantee the provision of feedback to participants in the most straightforward approach. Practitioners in the professional industry believe that when translators are asked to review the changes introduced in their translation by the reviser, this promotes awareness of the translation quality produced and allows the translator to take any feedback into account in future work (STP, N.D.). A similar situation occurs in crowdsourced projects: for those with strong motivations and the commitment they made when signed up to the project (*Gutenberg Project* in particular), giving and receiving feedback is a natural tendency. Most participants showed appreciation regarding the feedback provided by peers and external reviewers, which also matches the findings of the questionnaire survey in which 93% of the respondents demonstrated that they value others' feedback. Moreover, learning is also reflected in the discussion of translation problems via the chat board (for more details, see Chapter 6).

5.5 Summary

This chapter investigated the participation and organisation of crowdsourcing translation projects through workflow mapping. First of all, it confirmed the advantage in having the crowd process a high volume of translation with reduced financial cost. Participants in the crowdsourced environment play multiple roles in a project and they are able to spontaneously switch their roles according to the progress of the project. The chapter also verified the self-declared quality control methods in the questionnaire survey: self-revision, peer-revision, shared glossary list, and group discussion were all observed in the empirical data.

However, the widely-claimed advantage of ‘performing at high speed’ seems to be an over-estimation of crowdsourcing. The result of my **workflow mapping** indicated that revision takes much longer than translation for both less-governed and strictly-governed crowdsourced projects. My **productivity analysis** demonstrated that, indeed, the ‘translation’ speed for large volume text is high, but if quality is taken into consideration, too, the crowd needs more time than professionals to perfect their translation.

Secondly, two kinds of workflow were identified: a **user-driven flexible workflow** that mainly depended on the participants’ spontaneity, and a **waterfall workflow** managed and monitored by the PM. The former represents a **network organisational structure** characterised by a high level of autonomy, while the latter represents a **hierarchical organisational structure** in which less autonomy is found. Both types work in the crowdsourced environment, giving participants different levels of control in the two organisational structures.

Thirdly, the **self-assessment** conducted by the participants in the *Gutenberg Project* confirmed the presence of **a formative learning process** in crowdsourcing translation projects. For collaborative projects with TEP-like workflows, elaborate feedback provided during the revision stage was acknowledged to be useful in improving language and translation ability. Project management-related skills were also identified as part of the new skills acquired by some participants. Therefore, TEP-like workflows in the crowdsourced environment are an important factor that helps fulfil people’s extrinsic motivation to enrich their human capital. Participants also emphasised the value of group discussion in facilitating learning which is analysed in Chapter 6, while 6.4 examines the effectiveness of the workflows in the observed projects evaluated by means of a translation quality assessment.

Chapter 6 Findings of the observation stage: 2 – communication

The previous two chapters have revealed that both respondents in the questionnaire survey and participants in the *Gutenberg Project* highlighted the important role of communication (group discussion) in motivating their participation. Intense discussions were noticed throughout the five projects during my observation. This chapter investigates the phenomenon of on-line peer-to-peer interaction in collaborative translation projects in the crowdsourced environment. The analysis focuses on identifying the characteristics of interaction that contribute positively towards effective collaboration. The data used for this purpose were extracted from the chat board, and subsequently annotated and discussed according to the Dialogue Act Analysis (DAA) and Social Network Analysis (SNA) frameworks.

DAA is used to interpret the categories and frequency of communication based on the interaction typology presented in Table 4 in section 3.2.3.1, Methodology chapter. It examines the communication patterns in crowdsourcing translation projects and the peer-learning process embedded in communication. SNA is used to visualise the information flow and communication pattern. It also helps identify the relationship between the participants, which is a used as a metric to evaluate the participants' contribution. The findings are compared with the interactions that took place in the *Students' Team project* which replicates multilingual localisation projects in the professional industry (section 6.3). The aim of this analysis is to establish a comprehensive set of communication features in collaborative translation projects. Moreover, as there are few empirical studies that explore the interactive aspect in translation practices, the comparative study aims to fill the gap in communication studies in the translation domain.

6.1 Data description – interaction categories and dialogue acts

Communication data were collected through the chat board in the collaboration platform that my observed projects took place. Table 17 briefly outlines the participants' involvement in communication, and shows that not all participants (both primary and supplementary participants) in the projects were actively engaged in interacting with peers in the chat board.

Table 17. Participants' involvement in on-line peer-to-peer communication

Projects	Number of Participants	Number of participants in the chat board
Health Project I (HP I)	12	7
Health Project II (HP II)	9	8
Business Project I (BP I)	11	10
Business Project II (BP II)	6	4
Gutenberg Project (GP)	3	3

To interpret the content of interaction and participants' performance, I first retrieved the communication record of each project and then annotated dialogue acts by identifying the sender and receiver(s) of the message, and then categorising the message based on the interaction typology (see Table 4). Tables 18-22 present statistics of dialogue acts performed by individual participants in each project. For consistency consideration, just as in the workflow mapping (section 5.1), participants are referred to as T1, T2, […] T10 (Translator 1, Translator 2, [...] Translator 10) in line with their task allocation; PM represents the project manager, and R the third-party revisers (R1, R2, R3). For convenience, the abbreviation of the dialogue acts is expanded to full terms below each table (for explanations of the dialogue acts, please see Table 4).

Table 18. Interaction statistics in *Health Project I* (*HP I*)

Interaction Typology		Translators						Project Manager	Total	Sum
	Dialogue Acts	T1	T4	T6	T7	T8	T10	PM		
General	CM	-	1	1	-	-	-	6	8	24
	CT	-	-	2	-	1	2	1	6	
	SO	-	1	1	2	1	2	3	10	
Information	GIF	1	-	-	-	-	2	1	4	53
	IR	-	1	2	1	1	1	-	6	
	IC	-	1	-	1	-	-	6	8	
	TR	-	-	-	-	1	-	-	1	
	TC	-	-	5	-	1	4	-	10	
	XR	-	-	-	-	1	-	-	1	
	XC	-	-	4	-	1	4	-	9	
	RR	-	2	1	-	3	-	-	6	
	RC	-	-	4	1	-	3	-	8	
Maintenance	MT	-	-	-	-	-	-	1	1	7
	AK	-	-	-	1	1	3	1	6	
Status	SC	-	-	-	-	-	-	-	-	5
	SR	-	-	1	-	2	-	2	5	
	Total	1	6	21	6	13	21	21		**89**

(CM – commission, CT – comment, SO – social, GIF– general information, IR – information request, IC – information clarify, TR – term request, TC – term clarify, XR – text request, XC – text clarify, RR – revision request, RC – revision comment, MT – motivation, AK – acknowledgement, SC – status check, SR – status report)

Table 18 presents the dialogue acts performed by each participant (as a message sender) in *Health Project I*. *T2, T3, T5, T9* and *T11* never interacted with their team members. *T1* only appeared once in the chat board and it was to state the colour which represents her work (GIF). *T6, T10* and *PM* contributed the most to the conversational discourse in terms of the *quantity* of dialogue acts, accounting for 24% of each. Social interaction (SO) and TC (term clarify) are the most common dialogue acts (11% of each), followed by XC (10%).

Table 19. Interaction statistics in *Health Project II* (*HP II*)

Interaction Typology		Translators						Project Manager	Reviser	Total	Sum
	Dialogue Acts	T1	T2	T3	T5	T6	T7	PM	R		
General	CM	3	-	-	-	1	-	6	1	11	137
	CT	8	3	-	-	19	11	7	11	59	
	SO	11	-	-	-	23	7	14	12	67	
Information	GIF	2	-	-	-	1	2	3	-	8	115
	IR	1	-	1	-	7	2	3	3	17	
	IC	5	-	-	-	3	4	2	-	15	
	TR	2	-	1	-	1	3	-	6	13	
	TC	7	-	-	-	6	6	-	7	26	
	XR	1	-	-	-	4	1	-	3	9	
	XC	2	-	-	-	1	-	-	3	6	
	RR	-	-	-	1	1	-	-	2	4	
	RC	4	-	-	-	3	2	-	8	17	
Maintenance	MT	1	-	-	-	1	1	2	1	6	18
	AK	2	2	1	-	2	2	1	2	12	
Status	SC	1	-	-	1	1	1	1	1	6	25
	SR	6	2	-	1	2	2	3	3	19	
	Total	56	7	3	3	76	44	42	63		**295**

(CM – commission, CT – comment, SO – social, GIF– general information, IR – information request, IC – information clarify, TR – term request, TC – term clarify, XR – text request, XC – text clarify, RR – revision request, RC – revision comment, MT – motivation, AK – acknowledgement, SC – status check, SR – status report)

In the *Health Project II* (Table 19), *T6* is the one who made the *largest amount* of contributions to the group interaction, accounting for nearly 1/4 of all dialogue acts among the eight participants, followed by *R* and *T1*. *T4* withdrew from the project before performing any dialogue act. SO (social interaction – 23%) occupies the top position among all the dialogue acts, followed by CT (comment on the *project* and discussions of *translation-related* problems – 20%) and TC (term clarify – 9%).

Table 20. Interaction statistics in *Business Project I* (*BP I*)

Interaction Typology		Translators						Project Manager	Revisers			Total	Sum
	Dialogue Acts	T1	T2	T3	T4	T5	T6	PM	R1	R2	R3		
General	CM	1	5	-	-	-	1	11	2	3	2	25	249
	CT	1	25	2	3	5	15	30	12	15	5	113	
	SO	-	27	1	4	4	24	18	6	12	15	111	
Information	GIF	-	3	-	3	-	1	10	2	3	1	23	197
	IR	1	7	2	2	-	3	4	-	4	2	25	
	IC	-	1	-	1	-	2	10	3	4	3	24	
	TR	-	6	-	-	4	3	3	-	3	5	24	
	TC	-	9	-	2	2	5	9	5	6	-	38	
	XR	-	5	1	-	-	2	1	2	4	-	15	
	XC	-	1	-	1	-	3	3	1	4	-	13	
	RR	-	2	-	-	-	-	-	-	1	1	4	
	RC	-	6	-	-	-	2	8	3	6	6	31	
Maintenance	MT	-	-	-	-	2	2	4	-	2	-	10	38
	AK	-	2	1	1	-	11	6	3	2	2	28	
Status	SC	-	-	-	-	-	2	1	2	1	2	8	44
	SR	1	10	-	3	-	1	8	1	6	6	36	
	Total	4	109	7	20	17	77	126	42	76	50		**528**

(CM – commission, CT – comment, SO – social, GIF– general information, IR – information request, IC – information clarify, TR – term request, TC – term clarify, XR – text request, XC – text clarify, RR – revision request, RC – revision comment, MT – motivation, AK – acknowledgement, SC – status check, SR – status report)

In the *Business Project I* (Table 20), the PM and T2 stand out as the two most active participants in the group interaction, initiating around 1/4 and 1/5 respectively of all dialogue acts. CT and SO are the most frequent dialogue acts (21.4% and 21%, respectively). Because the topic of this project was very intriguing (according to the participants' comments in the chat board), it aroused extensive discussion among participants about *the context of the source text*. Also, there was a lessons-learned session organised by the PM at the end of the project, in which participants expressed feedback and

suggestions for the future organisation of collaborative translation projects in *Yeeyan,* which is reflected in the large number of CT in *BP I*.

Table 21. Interaction statistics in *Business Project II* (*BP II*)

Interaction Typology		Translators			Project Manager	Total	Sum
	Dialogue Acts	T1	T2	T3	PM		
General	CM	1	1	1	2	5	61
	CT	8	3	2	3	16	
	SO	21	11	7	1	40	
Information	GIF	4	1	1	-	6	47
	IR	4	-	2	2	8	
	IC	3	2	-	2	7	
	TR	2	-	-	-	2	
	TC	-	2	3	-	5	
	XR	2	-	1	-	3	
	XC	-	1	1	-	2	
	RR	5	1	-	-	6	
	RC	-	2	6	-	8	
Maintenance	MT	2	-	-	1	3	11
	AK	4	1	1	2	8	
Status	SC	2	-	1	2	5	21
	SR	10	1	4	1	16	
	Total	68	26	30	16		**140**

(CM – commission, CT – comment, SO – social, GIF– general information, IR – information request, IC – information clarify, TR – term request, TC – term clarify, XR – text request, XC – text clarify, RR – revision request, RC – revision comment, MT – motivation, AK – acknowledgement, SC – status check, SR – status report)

In the *Business Project II* (Table 21), the external reviser (R) did not engage in the interaction and only performed revision on the edit panel. T1 contributed the largest number of dialogue acts in peer-to-peer communication (48.6% of the total). And again, SO is the most frequent dialogue act (28.6%), followed by CT (11.4%) and SR (11.4%).

Table 22. Interaction statistics in the *Gutenberg Project* (*GP*)

Interaction Typology		Translators		Project Manager	Total	Sum
	Dialogue Acts	T1	T2	PM		
General	CM	2	3	35	40	202
	CT	54	41	47	142	
	SO	4	4	12	20	
Information	GIF	3	7	60	70	198
	IR	11	4	16	31	
	IC	8	11	14	33	
	TR	3	4	9	16	
	TC	7	7	11	25	
	XR	1	4	7	12	
	XC	5	0	6	11	
	RR	0	0	0	0	
	RC	0	0	0	0	
Maintenance	MT	0	0	7	7	135
	AK	43	49	36	128	
Status	SC	0	0	7	7	60
	SR	16	16	21	53	
	Total	157	150	288		**595**

(CM – commission, CT – comment, SO – social, GIF– general information, IR – information request, IC – information clarify, TR – term request, TC – term clarify, XR – text request, XC – text clarify, RR – revision request, RC – revision comment, MT – motivation, AK – acknowledgement, SC – status check, SR – status report)

Different from the four *General Collaborative Projects*, social interaction (SO) is not significantly present in the *Gutenberg Project* (it only accounted for 3.4% of all the dialogue acts), which may be because the group discussion sessions were scheduled formally by the PM for specific purposes, e.g. for task assignment, instruction delivery, and translation problem-solving. The conversations usually spread around the initial purpose, but were limited to the topic of the translation project per se. Moreover, since the participants did not join the *Gutenberg Project* for the pursuit of social relatedness (self-declared in the follow-up interview), less social interaction was performed. CT, AK, and GIF are the three most frequent dialogue acts (accounted for 23.8%, 21.5%, and 11.8%, respectively). CT focused on the discussion of

project-related issues and *translation* problems between the participants; AK was mainly performed to signal the *receipt* of the PM's instructions, as well as the *agreement* achieved for the solution to translation problems; and GIF mostly contained the instruction and requirements of the translation project. The PM contributed the most to the group interaction (48.4%), and the two translators' contribution is evenly distributed (around 26%).

In general, the interaction typology (Table 4) encompassed all the dialogue acts in the observed crowdsourcing projects. This section briefly presented the annotated interaction data. The following sections analyse the interactions in detail to examine the individuals' extent of engagement in group communication and to find out the implications of dialogue acts, as well as the power relationship between participants connected by these interactions.

6.2 Data analysis

This section analyses the core communication activities in collaborative translation projects in the crowdsourced environment. Clark and Schaefer (1989:257) defined 'discourse' as 'a sequence of utterances produced as the participants proceed turn by turn.' When people take part in conversations, they are contributing to the discourse. The analysis of the interaction between the participants in the observed projects can be seen as an examination of the discourse which is another form of collective activity. Teasley and Roschelle (1993) put forward the idea of building a 'Joint Problem Space' (JPS) for learning purposes as students were found to 'use language and action to overcome impasses in shared understanding and to coordinate their activity for mutually satisfactory results' (ibid:26). Inspired by these pioneering works in conversational analysis, the framework of examining the communication activities in the crowdsourced environment is a discourse analysis that deals with not only the content of peer-to-peer interaction, but also how the interaction promotes or inhibits knowledge-sharing within the group.

6.2.1 Participation – turn-taking and multiple roles

Studies have demonstrated that social grounding is one of the characteristics of effective collaboration (Soller, 2001). The criterion for social grounding is that the team (the contributor to the conversation and the partners) have a mutual understanding of meanings (i.e. the content of the conversation) (Clark and Schaefer, 1989), manifested through the turn-taking between individuals. Discourse units such as questioning, clarifying, acceptance and disagreements represent various specific discourse forms available for taking a conversational turn. The structure of turn-taking sequences is an indication of the degree to which mutual understanding is established (ibid). In collaborative projects, participants use role-swapping to help prompt the turn-taking behaviour that contributes to social grounding (Soller, 2001). Such a characteristic was also present in the projects I observed in my research.

Participants naturally took turns playing characteristic roles between dialogue threads to form a productive conversational discourse. For example, in *Health Project II*, T6 played three different roles during three consecutive dialogue threads: the annotation of the interaction record demonstrated that T6 played the role of questioner in one segment – asking for clarifications such as, '*when shall we finish the first draft translation?*', '*why do I keep disconnecting to the platform?*'; advisor in the next segment – making specific recommendations regarding how a requested term should be translated; and encourager in another thread when one participant indicated that the source text was too difficult and consequently T6 motivated the peer's participation.

Such natural role-switching during communication was a very common phenomenon I observed in my research, suggesting that, in the crowdsourced environment, because there are few restrictions to an individual's role, a participant is able to contribute to the project from different perspectives. The accumulation of role-switching behaviours maintains the social grounding of a collaborative project, and also extends an individual's value in the crowdsourced environment.

The field of conversational discourse tends to focus on two phases: the

presentation phase and the acceptance phase (Clark and Schaefer, 1989). The former refers to A *presenting* an utterance for B to consider. The latter refers to B accepting A's utterance by giving evidence that he/she *understands* what A means. Many turns occur in dialogue acts to achieve the transition between these two phases. For example, in the *Gutenberg Project*, T1 would reword and explain why he translated one sentence in a certain manner, while T2 and the PM would question and debate the accuracy of the translation, thus leading to the acceptance stage.

With regard to the individuals' engagement in communication, as mentioned in the previous section (6.1), not all participants joined the conversation with team members. However, this does not suggest that those people's contribution is not valued in the projects; most of them finished the tasks that were initially assigned to them in accordance with the project requirements. Secondly, those who contributed to the conversations, no matter what dialogue acts they performed, helped promote the progress of the project.

The following section discusses how sequences of specific dialogue acts were performed, and their implications.

6.2.2 Dialogue Acts Analysis

Pie charts in Figure 25 and Figure 26 present the percentage of interaction categories in each project, highlighting that '**general**' interaction (commission – CM, comment – CT, and social – SO) and '**information**' interaction (general information – GIF, information request – IR, information clarify – IC, term request – TR, term clarify – TC, text request – TR, text clarify – TC, revision request – RR, revision comment – RC) take up the majority of communication in the crowdsourced environment, whereas interactions regarding '*maintenance*' (acknowledgement – AK, motivation – MT) and '*status*' occur much less frequently. This result can be explained by the fact that the four *General Collaborative Projects* were mainly user-driven, and the PMs were less involved in devising and coordinating the peer-to-peer interaction.

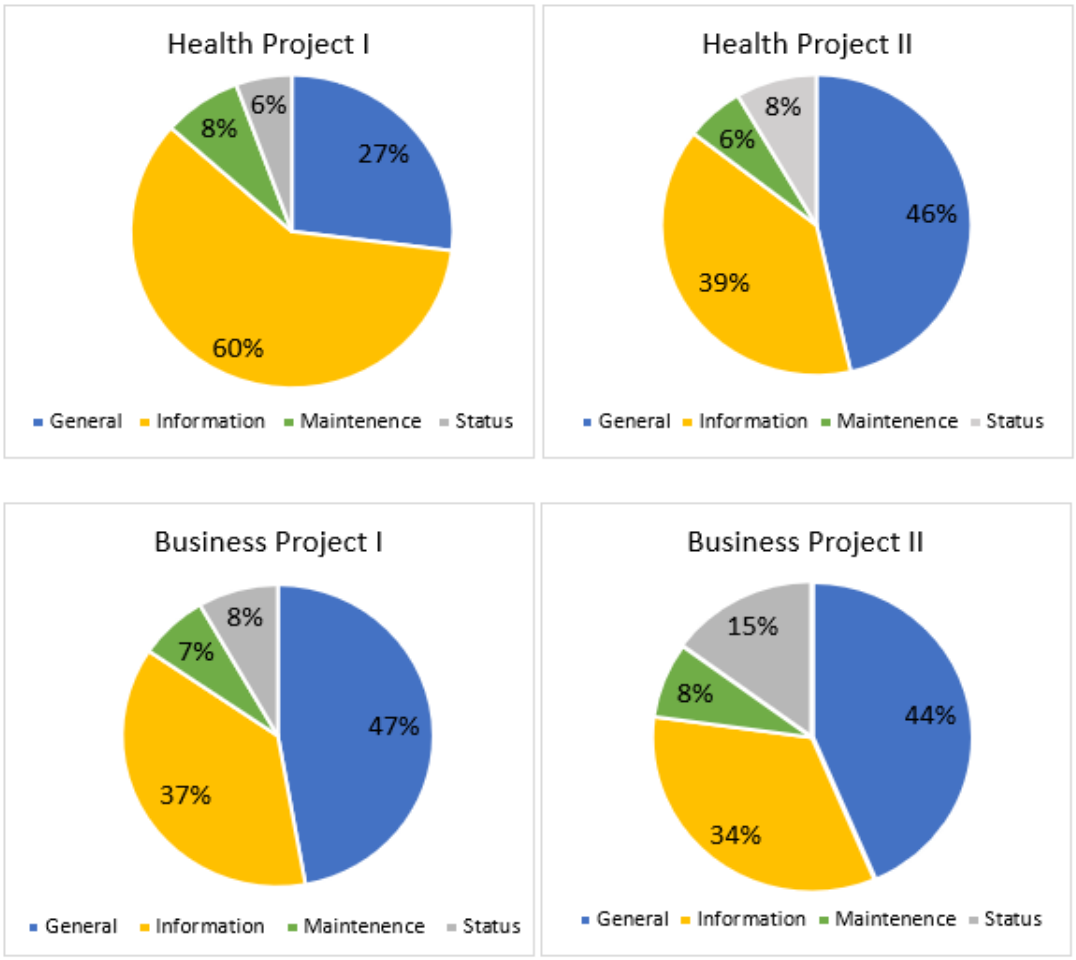

Figure 25. Percentage of interaction categories across the *General Collaborative Projects*

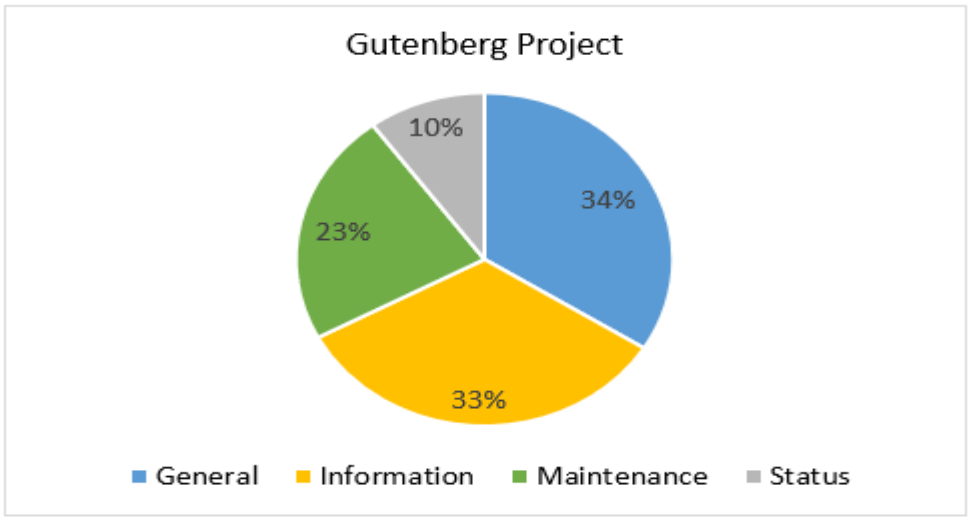

Figure 26. Percentage of interaction categories in the *Gutenberg Project* (*GP*)

However, in the *Gutenberg Project* (Figure 26), the percentage of '**maintenance**' interaction is higher than the four *General Collaborative Projects*, which is due to the hierarchical organisational structure of the project. The PM made great effort to organise an effective interaction between the participants. She proposed to save translation-related problems until the revision stage after the second round of revision (cross-check 2 and self-revision 2; see Figure 24). The PM also negotiated with the two translators the organisation of discussion sessions on two sequential nights. Since the discussions were formally organised, translation-related problems were discussed and agreed by all participants, which made the acceptance phase of the conversational discourse – 'acknowledgement' (AK) – appear frequently in *GP* (22%; see Figure 28).

Due to the different settings of the two kinds of crowdsourcing translation projects, there are several differences regarding the content and frequency of peer interactions. The following subsections analyse the dialogue acts in accordance with the interaction typology.

6.2.2.1 General interaction

First of all, in the '*general*' interaction category, '**social**' (SO) is the most common dialogue act in the four *General Collaborative Projects* (see Figure 27). Messages belonging to this category usually begin with '*Hi, XX, welcome to the project'*, '*Hi, here I am, anyone online?*', and end with '*I'm going to leave now, see you next time*', '*Good night*'. 'SO' can be used as an engagement indicator for individual participants. Though the topic of SO is not related to the substance of translation, it implies the participants' interest in collaborating with others in the task. In other words, if a participant constantly performs the dialogue act 'SO' with other participants he/she is still engaged with the project and is unlikely to forget about his/her task(s). Also, by following the distribution and frequency of the 'SO' dialogue act, the project manager or the organiser can detect the inactive participants in the project. If a translator does not engage in the interaction with peers or perform the task on the platform as others do, there is a risk that the translator may withdraw from the project early or delay the task. In such cases, it is better for the project manager to contact the inactive participant in

person to check the status of his/her task. Using SO as an indicator can be an effective way of keeping participants on track, as well as plan for unwanted developments, thus mitigating various levels of project risk.

As mentioned in section 6.1, SO was not frequently performed in the *Gutenberg Project* (see Figure 28) – in fact a major difference between the *General Collaborative Projects* and the *Gutenberg Project*. Due to the openness and autonomy in the former environment, in addition to translation-related activities, participants considered it as a platform to extend interpersonal communication. However, the latter was seen as a formal working environment in which the interactions focused on the substance of translation mainly.

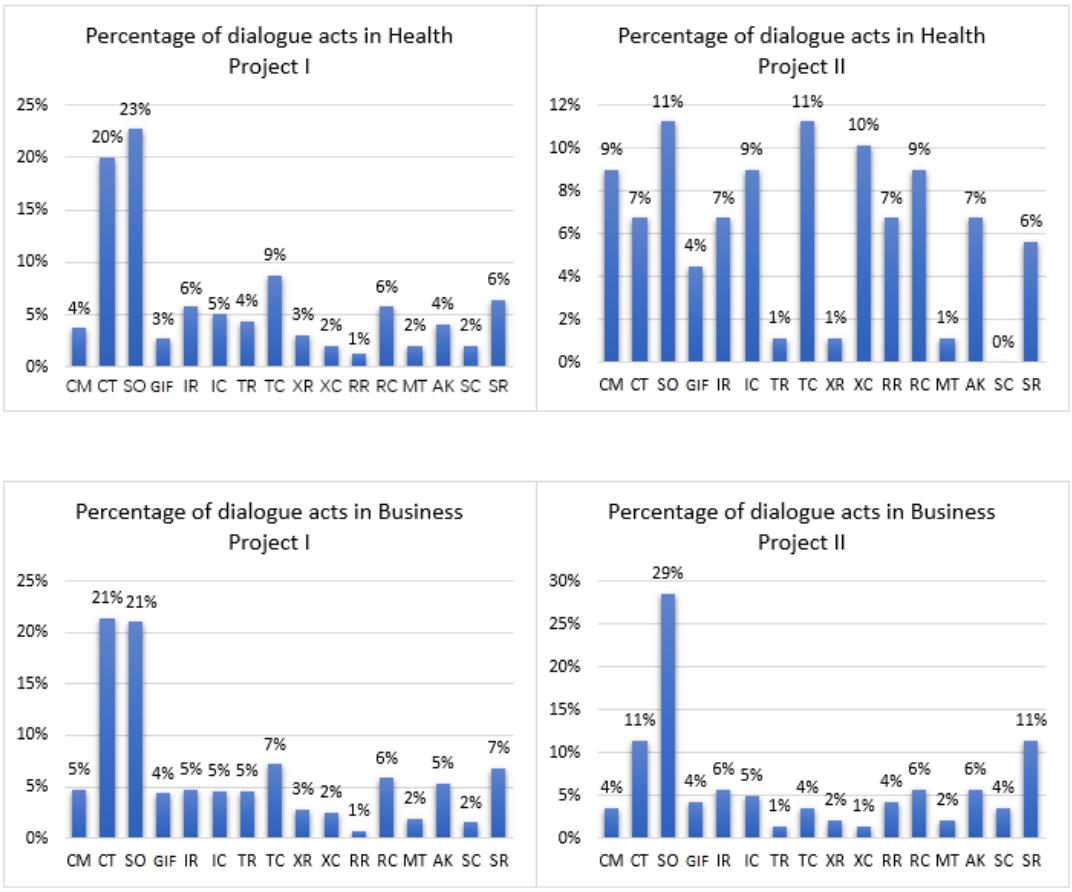

Figure 27. Percentage of dialogue acts in the *General Collaborative Projects*

The number of '**comment**' (CT) interactions almost reaches its highest among all dialogue acts in *Health Project I*, *Business Project I*, and the

Gutenberg Project, accounting for approximately 20% in each project. CT involved participants' comments on the project as a whole (i.e. source text, schedule, project instruction, and the collaboration platform). For instance, in *Business Project I*, T2 indicted that '*I think several people revising different parts of the text may lead to inconsistency*'; T6 expressed her opinion towards the edit pane of the collaboration platform, saying that '*The presentation format of having the source text at the top and the translation in the bottom is frustrating. I cannot get used to it'*. Such topics led to a continuous discussion in the chat board. CT interactions can therefore help the crowdsourcer find out what kind of topics the participants are interested in, and is a useful tool for collecting genuine feedback from participants, as it takes place in a natural environment. A close analysis of these interactions is beneficial for the future organisation and recruitment of participants in crowdsourcing translation projects.

Other topics in **CT** included participants' discussion on **questioning-related dialogue acts** (information request, term request, and text request). Because these dialogue segments were merely discussions without a direct answer, I categorised them as 'comments'. They can still be used as an indicator to see when an intervention should occur – for example, in the *Gutenberg Project*, the translators were not sure how to translate the name of a strait and they presented their own reasons for each of their translation proposals. However, no agreement was achieved after their discussion, leading the PM to propose consulting the professional editors in *Yeeyan*. The start of a 'comment' dialogue thread represents the *presentation* phase of the conversational discourse, while the sequential discussion represents the *acceptance* phase, meaning that the peers understand what one attempts to express in the previous utterance, and the exchange of opinions is the evidence of understanding.

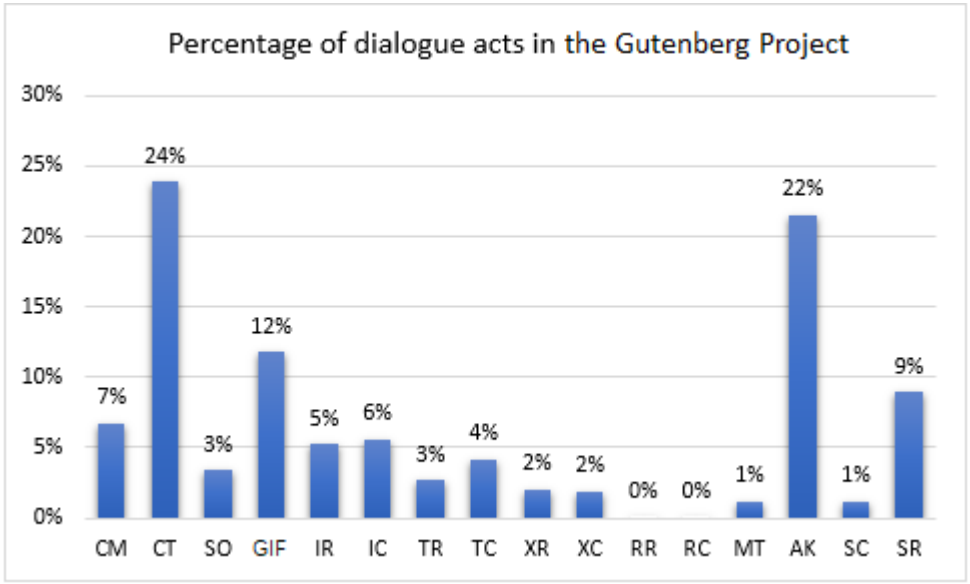

Figure 28. Percentage of dialogue acts in the *Gutenberg Project*

The relatively low proportion of '**commission**' (CM) in dialogue acts indicates that not all participants tend to clearly claim for a specific task through the chat board; instead, some started working within the edit pane straightaway. CMs performed by the PM were 'task assignment', which were a supplementary requirement in addition to the initial project instruction. As Figure 27 shows, the percentage of CM in *Health Project II* is higher than the rest of the *General Collaborative Projects*, explained by T4's withdrawal from the project without notification and without completing his/her work. As presented in the workflow mapping in Figure 21 in Chapter 5, on *Day 6*, the PM performed CM by asking if someone can translate T4's part. Moreover, the PM assigned the extra task and asked participants to perform peer-revision at the latter stage. Prior to the request, revision was mostly self-revision and participants were not intensively involved in revising others' translation.

Section 5.3.1 has already discussed the evidence suggesting that the crowdsourced environment for the *General Collaborative Projects* represents high autonomy, and the network organisational structure suggests a mobile personnel relationship between the PM and the participants. The discourse of task assignment performed by the PM is in line with the finding of the workflow analysis. The PM's task assignment through the dialogue act CM

did not allocate specific tasks to specific individuals. The utterance was presented in a very general way: '*can someone translate Part 4?*' and '*guys, could you please cross-check each other's work?*' Participants still had the freedom to decide whether to perform the extra task or not, which makes participating in crowdsourcing activities attractive for some. However, this much freedom brings several project management challenges: if participants give clear, unambiguous responses to task assignments, it is easier for the project manager or crowdsourcer to predict the schedule of the project and to monitor the progress of the claimed tasks. From this standpoint, dialogue acts regarding CM demonstrate the importance of a project manager's close involvement in the crowdsourced environment.

6.2.2.2 Information interactions

The previous section discussed the implications of dialogue acts from the perspective of project management – i.e. dialogue acts that help predict the development of the project. This part analyses information interactions from the participants' perspective. Conversational discourse in the *information* category evaluates the individual's performance and the competence of collective intelligence in crowdsourced projects as they are more closely related to the substance of translation.

First of all, the statistics presented in Figure 27 and Figure 28 indicate that the number of IR (information request) interactions nearly equals the number of IC (information clarify) ones in all observed projects. Moreover, the percentage of TC (**term clarify**) is higher than the percentage of TR (**term request**); XC (**text clarify**) is higher than XR (**text request**) in *Health Project II*, and XC nearly equals XR in the two *Business Projects* and the *Gutenberg Project*.

This suggests that requests are processed through peer-to-peer interaction. The fact that the number of TC exceeds the number of TR is because: 1. there were multiple responses for one term request; 2. some of the revisers preferred to explain the term and to clarify why revision was needed, which does not necessarily to be a response to TR.

Table 23. Examples of TR, TC and CT in *Health Project II*

14:50 T1: 发病率应该说增长吧？	*T1: Can '发病率' [morbidity] collocate with '增长' [increase]？*	*TR*
14:50 T1: 不对应该是升高	*T1: No, perhaps it should be '升高' [grow]*	*TC*
14:51 R: 增长。。、、好像均可	*R: '增长'… [increase] It seems both terms are ok.*	*CT*
14:52 T1: 感觉提高或者升高比较舒服	*T1: I feel '提高' [raise] or '升高' [grow] sounds more natural.*	*CT*
14:52 R: 比率一般说高低	*R: 比率 [rate] usually collocates with (高低) [high/low]*	*CT*
14:53 T1: 那就升高。。	*T1: Let's use '升高' [grow].*	*TC, AK*

Table 23 presents an excerpt of TR, TC and the related discussion (CT). T1 and R in *Health Project II* talked about a collocation problem; T1 made the initial term request (TR) and proposed translation options (TC). R expressed her opinion and presented the evidence for how the collocation works in Chinese (CT). T1 accepted R's opinion and a final decision was agreed (TC and AK).

When annotating the conversational discourse, I found that one segment can contain multiple dialogue acts. '**Comment**' (CT) and '**acknowledgement**' (AK) are two frequent acts that accompany request- and clarify-related dialogue acts. The former one represents provisional views or personal judgements regarding the matter raised; the latter represents the result of the interaction if it appears together with 'clarification' – i.e. announcing the accepted solution to group members. This data demonstrates how communication can help participants solve problems with collective intelligence. More importantly, it highlights the collaborative learning process in the observed projects. The process of requesting information, providing and receiving feedback enables participants to exchange knowledge of how and why translation should be done in a certain way.

My observation found that RR (**revision request**) usually took place when someone was not sure about the translation. The person made a request by asking someone else to help check the challenging part, e.g. a sentence or a paragraph. Alternatively, RR was also performed when someone found that a 'part' (*part* here refers to the unit of task allocation) of translation had been

left unchecked. As discussed in the workflow mapping, in the four *General Collaborative Projects*, there were no explicit instructions regarding who should perform which task for which part. RR is a representation of **self-organisation** in the crowdsourced environment. The participants themselves coordinated the task and the schedule to make sure that every part of the text was revised. Generally, RC (revision comment) does not necessarily equal the feedback towards a specific RR. I noticed that most of the dialogue acts regarding RC were an evaluation or suggestion regarding a person's translation on a general level in the observed projects, i.e. 'formative' or 'simple' feedback (see section 2.2.5 in the Literature Review). For example, one of the RC in *Health Project I* was '*I feel the translation in Line 9 is too verbose, please consider making it more precise, shorter and clearer*'. The dialogue act 'RC' is another learning opportunity in collaborative projects.

RR and RC was absent in the communication data in *Gutenberg Project*. Because it was under strict project management, the instructions were clear: to cross-check each other's work, which made it unnecessary to perform RR. Moreover, the participants did not make ambiguous comments to their peer's translation (RC); instead, they offered specific suggestions on how the translation should be improved through TC and XC.

Table 24. Examples of information interactions in *Health Project II* (RC, XR, XC, CT, AK)

22:24 T6: 再看看第六段黑体部分	*T6: Please help me check the bold sentence in paragraph 6.*	*RR*
22:31 R: 帕吉特质疑，是否存在一种导致癌症的物质，即所谓的致癌物。。？	*R: 帕吉特质疑，是否存在一种导致癌症的物质，即所谓的致癌物。。？ [Paget asked if there is 'one material for cancer, one carcinogen,' that 'may form different but closely allied compounds?']*	*IR*
22:31 T6: 两个 one, 我对 one 很困惑	*T6: There are two 'one's in the sentence. I'm confused.*	*XR*
22:31 R: 我觉得第二个 one carcinogen 是同位语	*R: I think the second 'one' is the appositive clause for 'carcinogen'.*	*XC*
22:31 T6: 噢	*T6: Right.*	*AK*
22:32 T6: 那下面的 one?	*T6: How about the 'one' in the following?*	*XR*
*22:32 R: 这是偶的**拙见。。。不知正确与否。。。***	*R: **It's just my humble opinion… I don't know whether it is correct or***	*CT*

	not.	
22:32 T6: 翻成一种我总觉得与原文相反……	*T6: If 'one' is translated as '一种'[category], I feel it expresses the opposite meaning of the source text.*	*CT*
22:32 T6: 嗯嗯 集思广益嘛	*T6: Eh, we can draw on collective wisdom and absorb useful opinions.*	*CT*
22:34 R: 什么叫没有'一个癌症'？听起来好恐怖	*R: What do you mean by '一个癌症' [one unit of cancer] [ST: We now know that there is no 'one cancer']? It sounds terrifying.*	*RC, TR*
*22:35 T6: 我就觉得不妥啊…… 可是翻成一种癌症**似乎也不好……***	*T6: I know it's improper…but '一种癌症' [a type of cancer] **seems problematic,** as well.*	*CT*
22:35 R: 是不是应该是'癌症种类繁多	*R: Maybe '癌症种类繁多'？ [a variety of cancer]*	*TC*
22:36 T6: 啊啊啊。。。	*T6: Ah…*	*AK*
22:37 R: 反正意思应该是，不止一种癌症。。。换句话说，就是种类繁多了？？？	*R: The meaning of the sentence should be, 'We now know there is more than one type of cancer'… In other words, it means a variety of cancer.*	*XC*
22:37 T6: 一语惊醒梦中人……	*T6: Your words awaken me…*	*AK*
22:37 R: material 翻译成原料是不是不妥？物质怎样？	*R: 'Material' translated as '原料' [raw material] seems improper, how about '物质' [substance]?*	*RC*
22:38 T6: 我曾经翻得就是物质 但因为我之前的假设就是错的 所以就全改错了	*T6: I translated it as '物质' [substance] before, but my assumption was wrong, my changes were wrong, as well.*	*CT*
22:39 R: 原料，好像一般是指工业中的生产原料。。。	*R: '原料' usually refers to the raw material in industrial production.*	*TC*
22:39 T6: 嗯 我的本意就是要把它改了……结果你那么一提醒……	*T6: Eh… My intention is to correct it… Now that you reminded me…*	*CT*
22:40 T6: 嗯嗯 我再看看…… 谢谢谢	*T6: I'll reconsider it…Thank you.*	*AK*
*22:40 R: 啊。。。。啊。。。**不好意思。。。可能我理解有误。。。***	*R: **Ah…Excuse me… Maybe I misunderstood [the meaning of the sentence].***	*CT*
22:40 T6: 你对的呢……	*T6: Your understanding is right…*	*AK*
*22:41 R: 还有那个'一次性诊断'**是不是也有问题呢？***	*R: Also, '一次性诊断' [one-time test] [ST: there is no 'one test' for carcinogens] **seems problematic, as well**?*	*RC*
22:41 T6: 是啊…… 全是 one 一个系列的问题, 所以我全黑体出来了	*T6: Yes… They are all the 'one' series problems. This is why I put them in bold.*	*CT*
22:41 R: 我觉得后面的内容介绍的是很多种诊断方法也。。。	*R: I think '很多种诊断方法' [a variety approaches of testing cancer] is also [problematic]…*	*RC*
22:41 T6: 看着特别别扭, 但是自己	*T6: It reads very awkward. But I can't*	*CT,*

又想不出来……	*think of a better one…*	*XR*
22:42 R: 我觉得意思应该是。。对致癌物的诊断也是多方面的	*R: I think the meaning is… '对致癌物的诊断也是多方面的' [test for carcinogens depends on multiple aspects]*	*XC*
22:43 T6: 哈哈 嗯嗯 我跟之前的 one 都一个道理	*T6: Haha, eh, it's the same with my previous 'one' problems.*	*CT*
*22:44 R: 我觉得就应该和上面是一样的。。。。**拙见。。。不要介意哦。。。***	*R: I think it is similar to the previous problems… **My humble opinion… Please don' mind it.***	*CT*
22:44 T6: 是啊 谢谢你了！	*T6: Yeah. Thank you!*	*AK*

Table 24 presents a series of consecutive conversations between T6 and R in *Health Project II* in which all information-related dialogue acts are involved. It contains the translator's confusion (RR), the reviser's explanation (RC), and the exchange of their opinions (CT). It demonstrates how grammatical and semantic problems were solved through interaction. As in the previous example of TR and TC in Table 23, it reinforces the finding that knowledge exchange (CT) is essential in the transition between request- and clarify-dialogues. Such interactions provide a channel for the giving and receiving of individual 'elaborate' feedback, meaning that the explanation and evidence of revision is accepted by the participants.

Moreover, **the utterances in themselves** are worth noticing, as well: when offering feedback to others' translations, some participants demonstrated a polite and modest attitude (see the underlined, bold statements in Table 24) – for example: '*I feel…may be problematic*' or '*…seems problematic*'. Instead of making absolute statements, hedge words were used to mitigate and soften the tone, which created a *negotiation* space between the information provider and receiver. This approach may be explained by the traditional Chinese culture of being modest, or it can also derive from the nature of the crowdsourced environment, where participants treat each other as peers rather than members of a hierarchy as in the professional industry or formal training environments. Knowledge exchanges are thus processed in a more casual and relaxed atmosphere in crowdsourcing translation projects.

Table 25. Examples of GIF, IR and IC in *Business Project I* and *Health Project II*

17:55 T2: 想问一下，翻译好了之后是发给哪位还是？	*T3: I want to ask, when the translation is finished, who shall I send my work to?*	*IR*
19:26 PM: 直接把译文贴在这个文件里，然后全体译者和我一起对全文进行审校	*PM: You can copy-paste your translation on the Edit Pane; all translators and I will do a review for the full text later.*	*IC*
19:50 PM: 这个文件可编辑，可以直接在上面翻译，只是有时网速会掉线	*PM: You can edit the text on this platform or directly translate on it. It's just that sometimes it disconnects.*	*GIF*

23:47 T6: T1? 你在翻译?	*T6: T1? Are your translating?*	*SC*
23:48 T6: 把你遇到的专有词翻法发到上面啊	*T6: Please add the specialised terms to the glossary list.*	*GIF*
23:48 T1: 好，知道了	*T1: Sure, I will.*	*AK*
23:49 T6: 很晚了， 加油啊	*T6: It's too late now. Keep it up.*	*MT*

Another important dialogue act in the *Information* category is GIF (**general information**), which refers to the general provision of project-related information (e.g. translation requirements, project schedule, guidance for using the collaborative platform). The difference between GIF and IC lies in the anticipated recipients of the message: IC is the response to IR, making it an individual-specific message, whereas GIF is a generic message that needs to be disseminated to all participants. Normally, in the professional industry, GIF is one of the main tasks for the project manager – i.e. delivering project instructions and updating any changes related to the project specification.

However, in the crowdsourced environment, a notable fact is that the project manager was not the only person who performed the act 'GIF'. Many participants contributed to GIF, as well (see Tables 18-22 in section 6.1). Table 25 presents an excerpt of GIF made by the project manager in *Business Project I* and T6 in *Health Project II*. At least two translators in each project performed the dialogue act 'GIF'. In *General Collaborative Projects*, the frequency of translators contributing to GIF is higher than the PM (except for *Business Project I*), which suggests that, even though the participants were not committed to the project in a legal sense, they showed concern and responsibility for the progress of the project. They constantly sent reminders

in the chat board and informed team members about the rules of participation. This represents direct evidence of the self-declared sense of responsibility stated by participants in the questionnaire survey.

In terms of the *Gutenberg Project*, 86% of the 'GIF' dialogue acts were performed by the project manager. As discussed in the previous chapter (Table 15 in section 5.3.1), the project manager's performance in *GP* mirrored closely what is required in the professional industry, especially regarding her contribution in the chat board delivering project requirements (deadlines, file naming rules, formatting instructions, guidelines for adding footnotes, translation and revision rules), posting regular reminders for task submissions, and delivering feedback.

On the other hand, although e-mail is the main working tool for communication in the professional industry (Matis, 2014), the interaction between the project manager and the participants still mainly includes sending tasks and files, confirming receipt of messages and files, status checks and status reports (ibid:169-170), and more. Due to the heavy workload in the professional industry, considering that one project manager often has several projects ongoing on at the same time, communication management through e-mail is often very challenging. Online (often distributed) peer-to-peer interaction through the form of chat board appears to be more convenient and flexible.

In short, the broad category of *Information* interaction covers the dialogue acts that can maximise the sharing of information and knowledge in crowdsourced projects.

6.2.2.3 Maintenance

Dialogue acts in the *Maintenance* category are less represented (under 8% in the four *General Collaborative Projects* and 23% in the *Gutenberg Project*) compared with other interaction categories (see Figure 25 and Figure 26). In the *Maintenance* category, **motivation** is the least common dialogue act (2% in *Health Project I*, 1% in the rest of the *General Collaborative Projects* and the *Gutenberg Project*) (see Figure 27 and Figure 28). A possible reason is that participants in the crowdsourced environment are volunteers, meaning

that they are self-motivated to contribute to the whole project. In addition, more than half of the respondents to my questionnaire survey stated that once they signed up to participate, they were committed to the project and owned the responsibility to fulfil the task (discussed in section 4.5.2 in Chapter 4).

In Translation Studies, motivation is a frequently-discussed topic in project management (Dunne, 2011). However, how to apply this skill in real practice remains unclear. In the professional world, project managers are usually overwhelmed with several projects at the same time and they spend most of their time on scheduling (McNab and Walters, 2016). Working on motivating freelancers is the last thing that a project manager would consider doing, because, in reality, the quality and productivity of linguists has a direct impact on the latter's income and reputation. Motivation mechanisms, either from the project manager or directly from the client, seem to be less important in the professional world. However, in the crowdsourced environment, as indicated in the questionnaire survey, extrinsic motivations outweigh intrinsic reasons, so giving 'verbal' encouragement during group interaction represents a very easy approach to inspire participation, especially for those who participate to gain peer recognition. Although it does not relate to financial rewards, a dialogue act belonging to the 'MT' category demonstrates the respect, value and recognition towards one's work.

Together with motivation (MT), **acknowledgement** (AK) is another dialogue act belonging to the *Maintenance* category. These two dialogue acts share a common meaning in the sense of mediating and promoting a member's participation. In the observed projects, I noted that a majority of AK was performed to show appreciation, to compliment peers' work, or to express gratitude of one's contribution. Such dialogue acts can be considered as a variant of motivation mechanisms under Participatory Culture meant 'to gain recognition within the community' and suggest that 'everyone's contribution is valued'. It is also a gesture of politeness, which is beneficial in *community building*. AK is an indispensable factor that maintains social network cohesion and social order, especially in an Oriental culture.

Table 26 presents the examples of dialogue act AK in various conditions. The first two examples illustrate how complimenting and showing gratitude are reflected through AK. R1, as an external reviser in *Business Project I*, expressed the appreciation towards the participants' work, which was meaningful to the translators in the form of 'simple feedback'. In *Health Project II*, because T2's translation was the most problematic one (see the error annotation in section 7.1, Chapter 7), it received negative comments from peers. T2's response was to thank them for their efforts in revising his work and he demonstrated his determination to improve both his language and translation skills.

Table 26. Examples of AK in various conditions

16:38 R1: 好。。。你们的译文我一直在欣赏呢 (Business Project I)	*R1: Sure…I always enjoy reading your translation.*	*AK – compliment*
09:22 T2: 谢谢你们能够耐下心帮我修改，你们说的这些也更加激励我来学习英语，学习翻译，我注册这个网站也就才两周时间，翻译还停留在直译的水平。让你们为难了，在这里我也学到了很多，开始慢慢上道。看你们的文字我也收获了很多，同时我也看到了差距，好像其他部分改动的不是很多。虽然你们说的让我有点受刺激，不过这也是好事，知难而进，希望各位前辈多多批评。 (Health Project II)	*T2: Thanks for revising my part. What you guys said about my translation motivates me to learn harder both English and translation. I only became a Yeeyan member two weeks ago. My translation is at the level of literal translation. I know it took lots of effort for you to revise it. I learned much from this collaborative project and now I am starting to know how translation works. I gained a lot from your comments and your translation. I saw the quality gap between my translation and yours. It seems that there weren't many changes made in other parts. Your comments made me quite frustrated, but it is a good thing. I'll keep improving despite the difficulties I may come across. Hope you guys can help me and offer suggestions.*	*AK – gratitude*
21: 30 PM: T1, 嗯，我同意你的那个翻译 (Gutenberg Project)	*PM: T1, I agree with your translation choice.*	*AK – approval*
16:50 T6: 大家别忘了把专有名词添加到统词表啊 (Health Project I)	*T6: Guys, don't forget to add the specialised terms to the glossary list.*	*GIF*
16:51 T10: 嗯，知道了, 等会加 (Health Project I)	*T10: Sure, got it. I'll do it later.*	*AK – receipt of information*

The last two examples of AK in Table 26 relate to the information flow within the group. Collaborative translation projects in the crowdsourced environment represent a high degree of collective effort. AK is performed as the ultimate *acceptance phase* of conversational discourse, to signal the approval or receipt of information in the previous dialogue. Each AK dialogue act which follows the discussion of translation-related problems (TC, XC, CT and RC) or the provision of project-related information (GIF, IC) has three meanings: 1. one has received the message; 2. one thinks the message is useful; and 3. the discussion has successfully solved the problem. AK is the second most common dialogue act in the *Gutenberg Project,* it contained all these three interpretations (see Figure 28), and it highlighted how, for a formal, strictly managed project, AK is essential in maintenance as it is a sign of moving forward to a new stage (e.g. time to start discussing a new translation problem).

In short, AK represents a *positive* response towards someone's contribution either to the conversational discourse or the translation project per se, and can be used in combination with MT to reinforce participation.

6.2.2.4 Status

The *Status* category influences the workflow and progress of the project. My observations found that SC (**status check**) was usually performed under three circumstances: 1. the project manager checked the progress of the contributors' work; 2. participants checked the progress of a peer's work; 3. participants questioned the progress of the project. Examples of SC and SR (**status report**) are listed below.

Table 27. Examples of SC and SR in *Health Project II*

22:48 T6: 校对得怎样了啊?	*T6: How is the revision progress?*	*SC*
22:49 T1: 我校了第二段的一部分, 其他的不知道	*T1: I've revised some of Part 2. I don't know how the remaining parts are.*	*SR*
01:03 R: 你在校哪一段?	*R: Which part are you revising now?*	*SC*
01:04 T6: T1 的部分	*T6: I'm on Part 1.*	*SR*
07:49 T6: 一、三、四部分我全不校对了一遍；二部分我校对了后半部分；七部分我校对了前半	*T6: I've revised Part 1, 3 and 4, as well as the second half of Part 2 and the first half of Part 7. Can someone help me check the underlined sentences? I'm too*	*SR, RR, SO*

部分。打下划线部分大家看看吧……太累了，我去吃饭吃了…… 大家早安！	*tired. I'm going to have some food now… Good morning, everyone!*	
07:50 T6: 五六部分没看 大家一起看看吧~	*T6: Part 5 and Part 6 are left unchecked. Let's do them together.*	*SR, CM*

When annotating the conversational discourse, I found that SR occurred more randomly compared with SC – it did not necessarily have to be one's response to SC. Participants spontaneously reported the status of their work, which explained the higher proportion of SR than SC in Figure 27 and Figure 28. On the one hand, SR states the progress of the project as a whole for all contributors. On the other hand, it can be used as a strategy to avoid repetitive work, particularly in the crowdsourced environment when no explicit instructions are implemented. The *Workflow mapping* visualised in the previous chapter has demonstrated that some parts of the translation were repeatedly revised, which may be both inefficient and time-consuming in a crowdsourcing translation project with a tight schedule. In such circumstances, it is important for the crowdsourcer to keep an eye on the dialogue acts SC and SR because participants tend to inform peers of their progress along with other dialogue acts (as indicated in the last two examples in Table 27) in a sense of 'for your information' rather than a very formal report. SR could be used as an indicator to find out what kinds of tasks are being or have been performed, by whom, and to guide the future direction of the project – whether to add more functional roles or to adjust the workflow.

6.2.3 Social Network Analysis

Social network analysis is widely used across disciplines, such as transport studies, social studies, computer science, and criminology. Generally, it is used to investigate the patterns of relationships that connect social actors. Visual representations of social networks are important for understanding the network data and conveying the result of the analysis through displaying nodes and connections in various layouts, colours, and node sizes. This section maps the relationships and information flows between the participants in the observed projects.

As shown in Figure 29, each node represents an individual (i.e. the participant) and the ties or edges represent the connections (i.e. interactions) between them. When annotating the data, the whole team of each project is seen as a network that contained individuals fulfilling the roles of translator, project manager, reviser and commentator. The interactions between them form a large or small directed network with characteristics that can be deduced from the visualisation: the size of the node reflects the position of the participant in each network (the bigger the node is, the more actively the participant has contributed to the conversational discourse, and vice-versa). Colours are used to differentiate the primary roles of the participants. The frequency of the interactions between two participants is reflected in the weight of each edge – the thicker and darker the edge is, the more interactions there are, vice versa.

Figure 29 visualises the interaction data in the observed projects. All actors involved in each project (see Table 14, Chapter 5) are portrayed in their networks. As the graph shows, there were overlapping actors in the networks corresponding to *Health Project II*, *Business Project I*, and *Business Project II*. These projects were organised by the same text editor (PM 1) in *Yeeyan*, and six participants were engaged in more than one project acting as translator, reviser, and commentator as stated on the label of the nodes. In the graph, ‘T’ stands for *translators* who initially signed up to the project (i.e. they are primary participants), usually played multiple roles, and naturally switched roles when needed. ‘R’ stands for *revisers* who initially did not sign up to the project, but joined at a later stage (i.e. they are supplementary participants). The task they performed was revision alone. ‘C’ stands for *commentators* who, similar to the ‘R’, were not involved in the translation project per se, but contributed to the chat board and performed information exchanges with others.

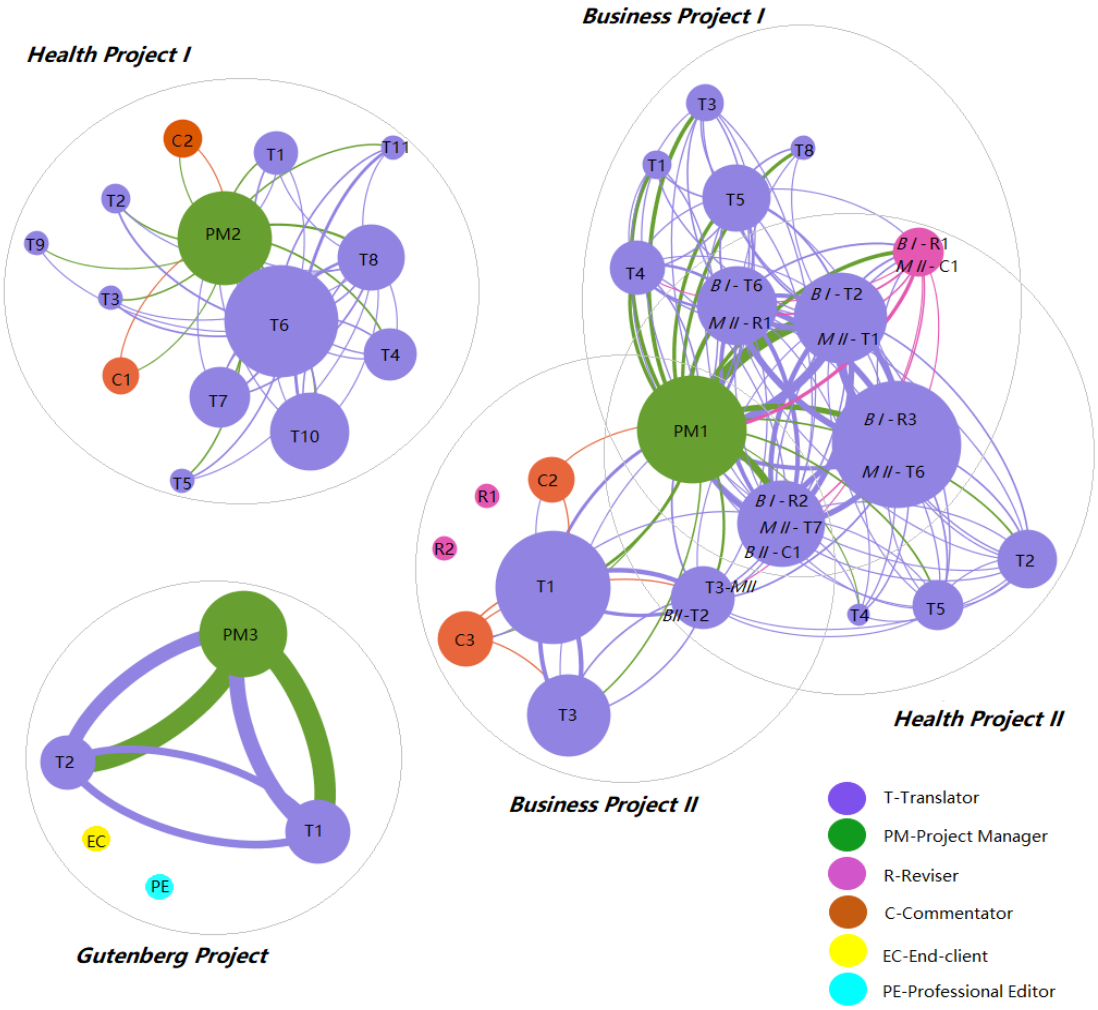

Figure 29. Visualisation of social network analysis of the observed projects

The most evident characteristic of social networks mapped in Figure 29 is the existence of large nodes in the network, representing **key actors** that contributed the most to the conversational discourse. According to the measurement of *weighted centrality out-degree*, in *Health Project I*, T6 contributed the most in the peer-to-peer interaction, followed by the project manager (PM2) and T10. They were the key actors in the interaction network. In *Health Project II*, T6 stood out to be the most active node, followed by PM1 and T1. In *Business Project I*, the project manager (PM1) appeared as the most prominent participant, followed by T2 and R2. In *Business Project II*, T1 occupied the central position of the network. The revisers (R1 and R2) silently performed the revision task one day before

finalisation, hence they had no connections with other participants in the network. Finally, the interactions in the *Gutenberg Project* network were led by the project manager (PM3).

The commentators (C) mainly performed two kinds of dialogue acts – IC and CT – by answering others' questions (e.g. on how to use the collaboration platform) and making comments on the project as a whole (e.g. translation progress). The smallest nodes in each project were the primary participants who did not engage in the interaction with peers (i.e. as message senders). They were considered passive message receivers in the networks and the connections to these nodes represent the messages being sent to them. To be more specific, they were T2, T3, T5, T9, T11 in *Health Project I*; T4 in *Health Project II*; and T8 in *Business Project I*. The isolated nodes in the *Gutenberg Project* were essential stakeholders (i.e. end-client and professional editor), however, their interaction with the translators and PM was not captured in the study.

Table 28. Attributes of key actors in each network

Networks	***HP I***		***HP II***		***BP I***		***BP II***		***GP***
Key Actors	T6	PM2	T6	PM1	PM1	T2	T1	T3	PM3
Weighted out-degree centrality	66	53	115	76	262	136	78	31	311
Lead participants (contributed the most in revision)	✓		✓			✓	✓		✓

Table 28 presents the result of the social network analysis of each project, focusing on the top two important actors. The identification of **key actors** in the interaction network is in line with the **lead participants** recognised through workflow mapping (section 5.1). It demonstrates that the 'active' participants are **more willing to** engage in a wide range of activities (in addition to their initially-assigned task), **investing more effort and time** in linguistic tasks (i.e. translation and revision) and **leading non-linguistic**

activity (e.g. peer-to-peer interaction).

Centrality measures are typically used as ‘indicators of power, influence, popularity and prestige’ (Scott and Carrington, 2011:4). For the network connected by interactions, the **weighted out-degree centrality** (for a detailed explanation of this measurement, please see 3.2.3.2, in the Methodology Chapter) can be interpreted as a form of *enthusiasm* (i.e. the willingness to provide meaningful information) and *competence* (i.e. the ability to help others solve problems). The following discusses how these ‘key’ actors influence the development of the project.

As described in section 6.1, the data shows that the frequent dialogue acts performed by ‘key’ actors were: SO (social), CT (comment), TC (term clarify), XC (text clarify), IC (information clarify), RC (revision comment) and SR (status report). First of all, SO is the most natural dialogue act in human interaction. In the four *General Collaborative Projects*, SO usually occurred at the beginning and end of most conversations. By initiating a conversation with SO, ‘key’ participants warmed up the atmosphere on the collaboration platform and stimulated others’ engagement, achieving a double effect: on the one hand, SO introduces more participants to join the conversation; on the other hand, it gives peers an impression that someone is working together with them. Following SO, other dialogue acts can be performed naturally and smoothly. CT represents the message exchanges regarding the project, and more importantly, it involves the discussion of translation-related issues. Similar to SO, CT introduces topics to the conversational discourse which discuss problems regarding word choice, difficulty in understanding the source text and syntactic rendering.

TC, XC, IC and RC are reactions to others’ messages, showing a constant information flow within the network – someone requests or questions, others respond. Through TC, XC and IC, ‘key’ participants in the observed projects acted as problem-solvers who offered direct solutions to peers’ problems, and stood out as major information providers in the networks. These dialogue acts contain suggestions that essentially improve the quality of the crowd’s translation.

SR is the most direct indicator for the progress of the project. My observation noted that the ‘key’ actors usually took the initiative of reporting the status of their tasks and, as a conformity effect, other participants followed their example both in terms of dialogue acts used and the behaviour displayed. In this way, the key actors further coordinated the crowd’s participation and rearranged tasks accordingly.

To summarise, in crowdsourcing projects, the key actors are those who contribute mainly as activators, information-providers, problem-solvers, and coordinators, who optimise the collaborative model and enhance the degree of knowledge exchange.

Another characteristic which emerged from the social network analysis is the interaction pattern in the crowdsourced environment. In terms of the interaction direction and the identity of interlocutors, I noticed multiple forms of interaction: ‘one-to-many’, ‘many-to-one’, and ‘many-to-many’. As shown in Figure 29, connections occurred not just between the key actors and the surrounding nodes, but also between the general participants, which created a ‘star-shaped network’ in each project characterised by the openness of crowdsourced environment in which participants are empowered with the channel and opportunity to establish social relationships. In the observed projects, because the interactions took place in the chat board, information flow was publicly available to all actors involved. High transparency in such interaction patterns widens the coverage of conversational discourse by, first of all, maintaining the individual’s autonomy in participation (i.e. freedom to decide whether or not to interact, with whom, and about what content), and secondly by providing unconstrained opportunities for participants to achieve reciprocity and self-efficacy by freely questioning, answering, and debating with a wide selection of conversation partners.

Table 29. Attributes of the network in each project

Networks	***HP I***	***HP II***	***BP I***	***BP II***	***GP***
Density	0.204	0.633	0.545	0.500	1
Percentage of translation errors corrected through interaction	3%	8%	13%	17%	4%

The density value of each network is a statistical proof of the presence of diverse forms in the interaction pattern: the closer the value is to 1, the denser the network is. A density value of 1 represents a *complete* network in which all actors are connected to one another – i.e. all the possible relationships in a network have been realised**.** Table 29 presents the attributes of the individual project networks based on the measurement of density and the percentage of translation errors corrected through interactions.

The *Gutenberg Project (GP)* is the only one that reached a complete interaction network. As the workflow and organisational structure of *GP* resembled a collaborative translation project in the professional industry, a simple top-down interaction model was expected. However, because the project was performed via the internet, under the PM's organisation, the participants overcame the geographical obstacle and gathered regularly for group communication regarding project-related and translation-related issues. This achieved a balanced relationship in which the visibility of information was maximised. According to Hanneman (2005), in a directed network, the extent to which a population is characterised by 'reciprocated' connections (those where each directs a connection to each other) may inform the degree of cohesion, trust, and social capital present. The density value of *GP* verified the reciprocal relationship between the actors, which was mirrored in the presentation and acceptance of utterances.

The interaction network of the *General Collaborative Projects* is sparser than that of *GP*. Without formal organisation, conversations were casual and less centralised. Moreover, the number of participants influences the density of the interaction network, as well: the more actors there are, the harder it is to

realise completed ‘reciprocated’ connections in the many-to-many interaction pattern.

As discussed in the previous section (6.2.2), the questioning, answering, discussing, and agreeing dialogue acts solved translation-specific problems. My quality assessment which I will present in next chapter (section 7.1) identified a variety of errors during different stages of the observed projects. To examine the influence of peer-to-peer interaction on improving translation quality, I used the frequency of error categories identified in the first draft translation as a parameter, and I calculated the percentage of errors corrected during subsequent network interactions. The correction rates through interaction were: 3%, 8%, 13%, 7%, and 4% in *HP I*, *HP II*, *BP I*, *BP II*, and *GP*, respectively. Lexical errors were corrected more frequently through interaction than other error types, which correlates with the relatively larger proportion of TC – term clarify dialogue acts presented in Figure 27 and Figure 28.

I also investigated the relationship between the density of the network and the effectiveness of the peer-to-peer interaction for improving translation quality. However, no significant relationship was found: due to the diversity in the content of interaction (discussed in section 6.2.2) and the features of the source texts, participants only turned to the chat board when they could not solve a translation problem on their own.

To sum up, through the social network analysis of the interactions between participants, I identified the existence of ‘key’ actors in crowdsourcing translation projects and examined the dialogue acts that set them in these key positions. Moreover, diverse interaction patterns (one-to-many, many-to-one, and many-to-many) were discovered, which promoted the development of a collaborative project and expanded the coverage of knowledge-sharing within the group.

The next section compares the result of the communication analysis from the *Yeeyan* crowdsourcing projects with a *Students’ Team Project* in an MA class in order to establish a comprehensive set of communication features in collaborative translation projects.

6.3 A comparison of communication analysis between *Yeeyan* crowdsourcing projects and University of Leeds MA in Applied Translation Studies *Students' Team Project*

6.3.1 Research Background

I conducted an archive research of the communication data in one *Students' Team Project* based on *dialogue act analysis* and *social network analysis*. The data was collected from one of the projects conducted during the University of Leeds *MA in Applied Translation Studies* programme. Students are expected to sharpen their skills on both theoretical and practical levels to meet the industry requirements, which is why, in addition to translation theory and specialised translation classes, the programme also focuses on providing students with training in the different facets of the Language Services Industry – from project management to working with Computer-Assisted Translation (CAT) tools, communicating effectively, and developing industry-relevant finance and marketing skills. In this context, the core CAT module provides specialised training in a variety of software and real-life practices seen in localisation projects.

During the CAT module, the instructor organises mock localisation team projects to allow students to experience a comprehensive range of complex challenges in realistic scenarios, including communicating across language barriers as well as time zones, planning and organising a translation/localisation project, collaborating with team members, and managing the dynamic relationship between freelancers, the language service provider and the end-client.

Generally, students are divided into teams and are assigned a localisation project centred on a pragmatic text in a variety of file formats (*HTML, PDF, Microsoft Office*) authored by various non-governmental organisations (NGOs). The instructor assumes the role of the client, while one of the students in each team acts as a project manager (PM), and the rest are

freelancers. In addition to supplying the source files, the instructor acts as a client and provides a project brief that includes a list of requirements, final deliverables, a schedule of deadlines and the criteria for evaluation. The project is then entrusted to the students to perform. There are four stages in the localisation projects:

- Stage 1: setup – the PM receives the project enquiry and confirms receipt to the client. The PM assesses the client's requirements, negotiates unclear elements with the client, then plans the schedule and sets up the project in a chosen CAT tool. The PM also contacts the freelancers, who quote for the language services they offer. Once the resources are all confirmed, the PM quotes the client and the client confirms the quote or raises additional questions beforehand. After the client's confirmation, the PM confirms the work to freelancers, and sends them the project brief and documents.
- Stage 2: translation/localisation – freelancers translate/subtitle/localise, and inform the PM when their task is ready (unless the PM is using a CAT tool which tracks the freelancers' progress automatically).
- Stage 3: revision – the PM informs the revisers, who check the translations and make comments. The revised translation is sent to the PM.
- Stage 4: finalisation – the PM checks that the localisation is complete, and that the quality control checklist has been filled in. Each student submits the localised file, the accompanying translation memory, a terminology database, and a personal assessment of the project. The final product is then delivered. Freelancers and the PM invoice their respective clients.

The instructor evaluates the students' performance in the team projects as part of their assessment in the CAT module. The students are required to communicate with each other via email, and to copy the email to the instructor for evaluation. In addition to email communication, there is a forum for students to raise questions and suggest answers to translation-related problems that they encounter in the projects.

To make a comparison with the crowdsourced environment and identify the key elements that contribute to the success of collaborative localisation projects, I investigated the students' communication data using the same methods I had used to analyse the crowdsourcing projects. The *Students' Team Project* investigated in this study had the following features:

- it was a multilingual localisation project
- it had 10 Teams (and 10 PMs)
- it also had 67 Freelancers: translators, subtitlers, DTP specialists, revisers, and terminologists
- the location of freelancers was in: Texas, US; Ljubljana, Slovenia; and Leeds, UK
- the language combinations of the students working as freelancers were: EN-AR, EN-ZH, EN-EL, EN-PT, EN-SL, FR-EN, DE-EN, IT-EN, JA-EN, RU-EN, and ES-EN[21]
- the tasks included: translation, subtitling, webpage and image localisation.

A new interaction typology (Figure 30) based on the one used in the crowdsourced environment was customised to fit the 'formal' activities which may be performed in the professional Languages Services Industry on which the Leeds Student Projects are based (e.g. *Recruitment* and *Financial acts*).

[21] The language codes represent: EN – English, ZH – Chinese, EL – Greek, PT – Portuguese, SL – Slovenian, FR – French, DE – German, IT – Italian, JA – Japanese, RU – Russian, and ES – Spanish.

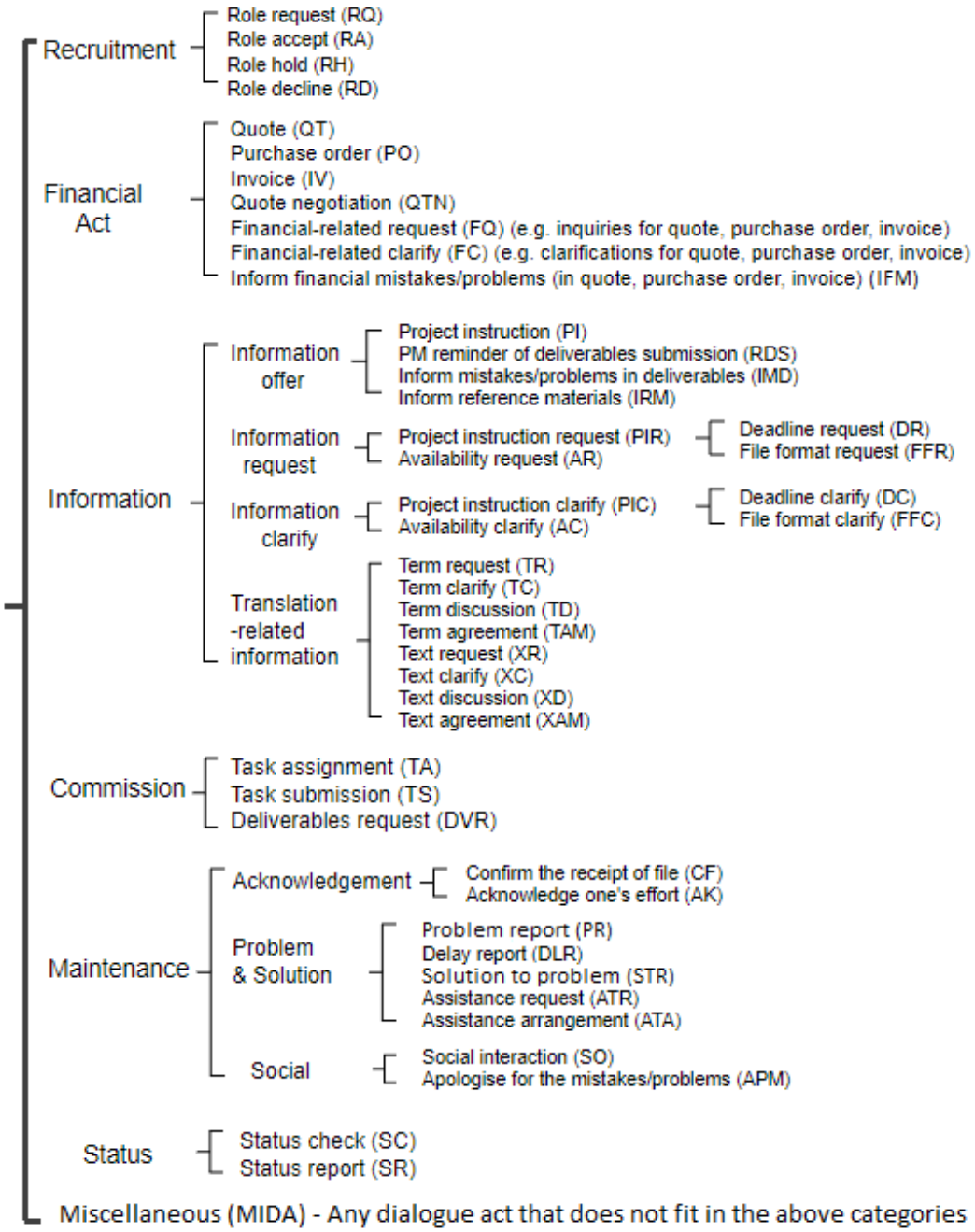

Figure 30. Interaction typology in the *Students' Team Project*[22]

[22] The interaction typology in this localisation project is derived from the interaction typology in the crowdsourced environment (see Table 4 in the Methodology Chapter).

6.3.2 Discussion

6.3.2.1 General findings

In this study, I analysed 2,295 student email exchanges and 126 forum posts. Each email contained 1-4 dialogue acts and all messages were written in English except for the Slovenian translators using their working language (Slovenian) in forum posts. Key points (e.g. deadline, file format) were presented in bold in the emails sent by the PMs, all of whom used an agreed email signature, indicating the (mock) agency name, contact information and their availability. Many freelancers also developed their unique email signature listing the services that they offer, as well as their contact information. Such details demonstrate the students' attitudes toward the team project: treating it formally and seriously in accordance with the professional conduct in the Language Services Industry.

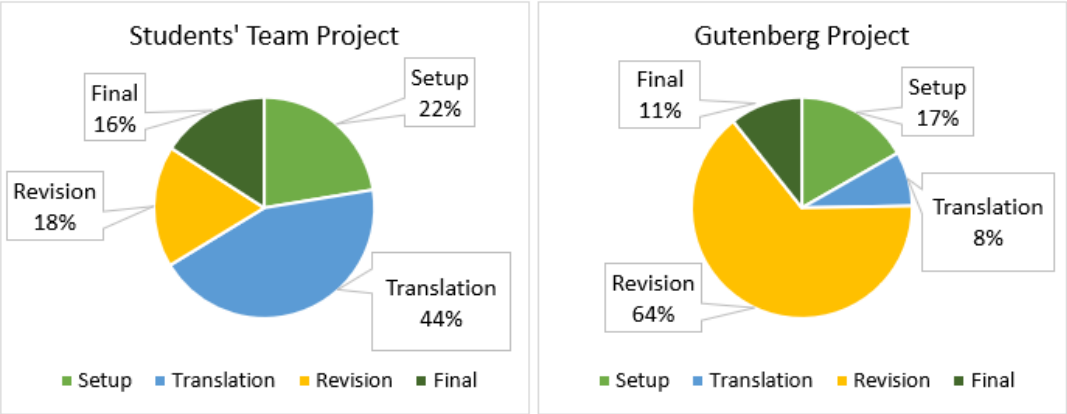

Figure 31. Percentage of interactions at different stages of the *Students' Team Project* and the *Gutenberg Project*

Similar to the *Gutenberg Project* in the crowdsourced environment, the *Students' Team Project* had an explicit schedule arrangement and a precise workflow. The interaction data was firstly analysed along the key stages of the project. As shown in Figure 31, nearly half of the interactions took place in the translation (localisation) stage (44%), followed by the setup, revision and final stages (22%, 18%, and 16%, respectively). By contrast, in the *Gutenberg Project*, the revision stage was where most of the interactions occurred (64%), followed by the setup, final and translation stages.

The *General Collaborative Projects*, on the other hand, were user-driven, which largely depends on self-organisation, and there was no clear division between each project stage. The number of interactions was calculated based on the workflow mapping and the timestamp of messages to *roughly* investigate the frequency of interactions during different stages (Figure 32).

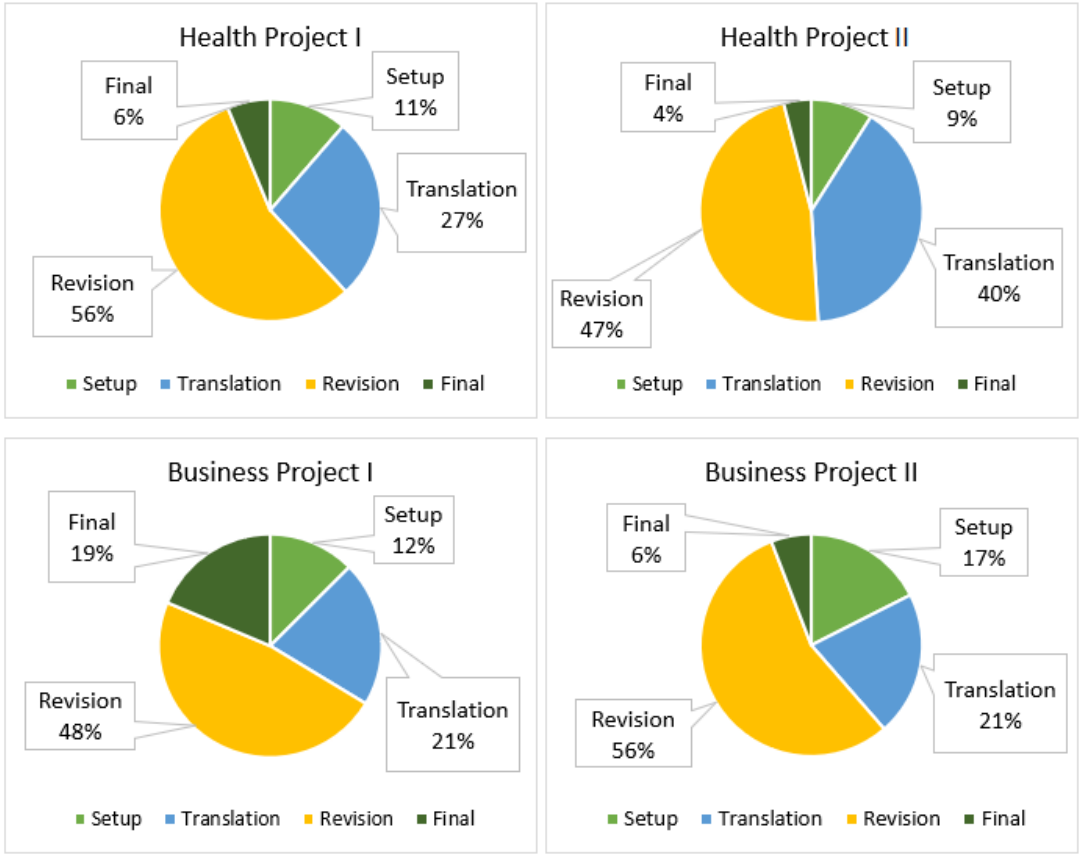

Figure 32. Percentage of interactions at different stages of the *General Collaborative Projects*

Again, by contrast with the *Students' Team Project*, around half of the interactions took place in the revision stage of the crowdsourced projects. The result corresponds to the workflow mapping in which revision appeared to be more significant, as well as its more influential for improving translation quality (see 6.4). For crowdsourced projects, it is in the revision stage that the quality of translation is greatly improved thanks to the help of peer revision. However, in the *Students' Team Project*, all translation-related problems were discussed and solved in the translation stage via forum posts.

6.3.2.2 Dialogue Acts Analysis

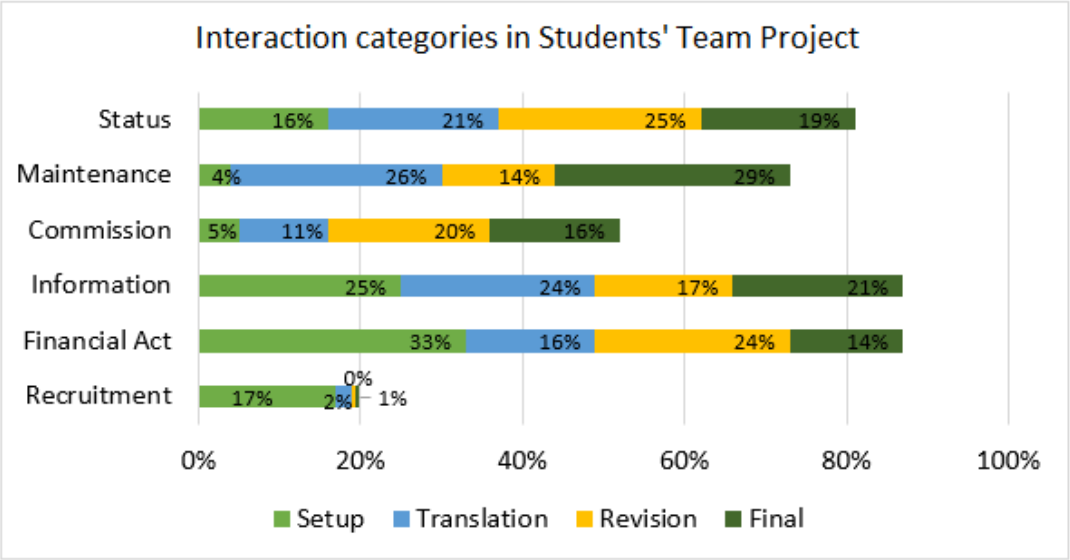

Figure 33. Percentage of interaction categories in the *Students' Team Project* with reference to the key stages

As Figure 33 indicates, the proportion of interaction categories varies as the *Students' Team Project* evolves. In the setup stage, *Recruitment* and *Financial Act* took up half of the interactions. In the translation stage, *Maintenance* (26%) and *Information* (24%) interactions were most prominent. In the revision stage, *Information* interactions decreased, while *Financial Act* and *Commission* increased by 8% and 9% respectively, compared with the previous translation stage. In the final stage, *Maintenance* was the most significant interaction category, followed by *Information*. The sections below elaborate on the importance of these interactions, and discuss their implications.

6.3.2.2.1 Recruitment and Financial Acts

Financial act occupied 22% of the interactions, in which quote (QT), quote negotiation (QTN) and other quote-related issues were the common topics. Financial act took up 1/3 of interactions in the setup stage of the *Students' Team Project*, just as one would see in the beginning of most professional projects.

In the emails for *Recruitment*, the PMs first introduced the required role and the source file (e.g. text type, length of the ST), then inquired about the freelancer's availability and rates, as well as other services that the freelancer can offer. The PM's availability was usually attached in the email for the convenience of future contact. Recruitment was the first step in the project, and student PMs usually spent 3-5 days to recruit freelancers for different services under the dialogue act 'role request' (RQ). The responding acts were 'role hold' (RD) (e.g. '*I'm interested in the project and I'll get back to you with my rates later.*') and 'role accept' (RA) (i.e. the PM and the freelancer reached an agreement with the proposed rates and task). In order to progress to the stage of 'role accept' (RA), a series of quote negotiations (QTN) were observed. The majority of QTN were initiated by the PMs asking some freelancers to lower their rates. However, there were also occasions in which QTN occurred when the PM actively proposed to raise the rate and asked for high quality work in return.

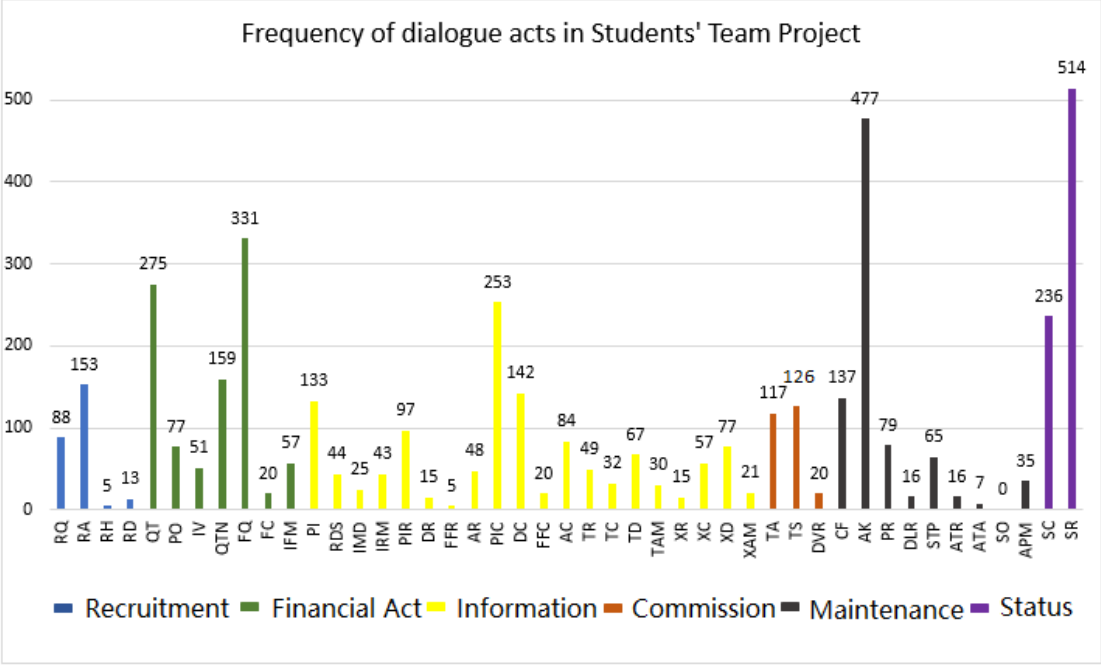

Figure 34. Frequency of dialogue acts in *Students' Team Project*

As shown in Figure 34, 'quote' (QT) is the fourth most frequent interaction among 45 kinds of dialogue acts throughout the project, explained by the following reasons: 1. There were two rounds of quotes throughout the project

(in the setup stage, *quote* refers to the *provisional* rate for the requested service; in the translation stage, after the freelancer received the task file, an *official* quote file was required). They were the subsequent response to FQ (financial-related request); 2. Some of the official quotes had infelicities (e.g. wrong file format; information missing) or serious mistakes (e.g. error in word count or price) that needed revision, which resulted the act IFM (inform financial problems/mistakes). Updated quotes were subsequently re-submitted; 3. Some of the PMs requested separate quotes for different kinds of services rather than a combined quote in the same project. Other financial acts included *invoice* (IV), *purchase order* (PO) and issues relating to the two subjects (IFM, FQ – *financial-related requests and inquiries*, and FC – *financial-related clarifications*).

However, in crowdsourcing translation projects, *Recruitment* and *Financial acts* were not observed due to the fact that the participations were voluntary. Yet, there was a similar dialogue act – 'commission' – in the crowdsourced environment which resembles the role of '*recruitment*' in the *Students' Team Project*. For the former one, the PM distributed tasks to all participants and the participants voluntarily signed up for the tasks; for the latter one, the PM enquires about the freelancers' availability and willingness to work, and the freelancers reply. The properties of the crowdsourced environment make the setup stage much simpler as opposed to a professional-like environment. The PM's main tasks were to initiate the project, equally divide the source text in accordance with the number of participants, and occasionally set rules for the participants to follow (e.g. specify the deadline and ask linguists to use the shared glossary list). However, these tasks were effective only when the crowdsourced environment had a large pool of active members enthusiastic about the translation and/or the topic of the source text. In the *Students' Team Project*, on the other hand, students naturally took on freelancer roles, following professional codes and ethics in the Language Services Industry. Therefore, the dialogue acts they performed were meant to fulfil the designated roles in a formal and professional sense.

6.3.2.2.2 Status

SR (status report) is the most frequent dialogue act in the *Students' Team Project* (see Figure 34). Similar to the result in the crowdsourcing projects, the number of SR is more than double that of SC (status check) (514 and 236 occurrences, respectively). SC was usually performed by the PMs under three circumstances: 1. the PM checked the progress of overseas freelancers, because their interaction was not as frequent as those in the UK; 2. the PM took the initiative to check progress when a freelancer did not reply to emails promptly; 3. the PM checked the progress when the deadline was close, but one had not yet submitted the task.

Status report occurred at a more random rate compared with *status check*. It not only included the freelancers' report on their progress, but also the PM's notification to freelancers, such as when a task file will be ready. When annotating the communication data in the *Students' Team Project*, I found that it was not necessary for someone to perform a SR only when the PM checked '*how are you doing with your work?*'. Freelancers usually informed the PM of their progress along with other dialogue acts in a sense of '*for your information*' rather than a formal report. The overall frequency of the *Status* interactions in the *Students' Team Project* and crowdsourcing projects is very similar, which is understandable given that keeping peers updated is a natural reflex in interpersonal interaction.

6.3.2.2.3 Information

Information interactions are divided into four subcategories in the *Students' Team Project*: *information offer*, *information request*, *information clarify,* and *translation-related information*. I found that the first one – *information offer* – was usually initiated by the PMs and it involved the following dialogue acts: project instruction (PI), PM reminder of deliverable submission (RDS), inform of a mistake/problem in the deliverables (IMD), and inform about reference materials (IRM). A notable fact is that only the PM in *Team 10* offered project instruction in the body of an email; the other PMs chose to merely inform freelancers that the instruction file was available in a shared online drive without further explanation. All PMs had sent a gentle reminder to the

freelancers with a specified deadline for task submission. The common ‘problems’ in deliverables were: *wrong file format*, *file cannot be opened*, and *missing deliverable*.

Information request included inquiries about project instructions (PIR), and availability requests (AR). *Information clarify* is the response to the former act. Similar to the crowdsourced environment, the number of information clarify acts also exceeds the information requests (499 and 165, respectively), which is due to the fact that in the *Students’ Team Project* a single request had several replies because the PMs needed to do further research (e.g. by consulting the instructor, peers and other resources) and re-specify the answer.

Moreover, deadline clarify (DC – 142 occurrences) was the second most frequent dialogue act among all information interactions, present throughout the project, and accounting for 28% of the *information clarify* category (see Figure 34). Given that there were few inquiries about the deadline (DR – 15), the PMs repeatedly emphasised ‘deadline’ (DC) in their email exchanges accompanied by the dialogue acts of role request (RQ) and task assignment (TA) to ensure that the deadline fits the freelancers’ schedules and the desired task can be completed on time.

Moreover, the content of information request in the *Students’ Team Project* is different from the crowdsourcing projects. Inquires in the team project were mainly on project requirements (e.g. instructions on localising various files, and desired file formats), whereas in the crowdsourced environment inquiries focused on the use of the collaboration platform and the deadline.

Translation-related information was separated from other subcategories in information interaction, which is different from the interaction typology used for crowdsourcing projects (see Table 4 in the Methodology Chapter). On the one hand, it is because these interactions were collected from forum posts (*ProBoard*) instead of email exchanges; on the other hand, discussions on translation problems are essential in the training environment for collaborative learning. Term-related interactions (term request – TR, term clarify – TC, term discussion – TD, term agreement – TAM) and text-related

interactions (text request – XR, text clarify – XC, text discussion – XD, text agreement – XAM) took up 4% each, which was less than in the crowdsourcing projects, considering the large number of multilingual files included in the *Students' Team Project*. It might be because the trainee freelancers' translation competence is higher than that of the participants in the crowdsourced environment (the localised files in the *Students' Team Project* were assessed by the instructor and some of them were successfully accepted and published by the end-client, i.e. non-governmental organisations).

6.3.2.2.4 Commission

The *Commission* category in the *Students' Team Project* included task assignment (TA – 117 occurrences), task submission (TS – 126), and deliverable request (DVR – 20) (see Figure 34). TA was often accompanied by two other dialogue acts – project instruction (PI) and deadline clarify (DC). DVR (20) took place only when the task was submitted and some of the requested deliverables were missing. It is similar to the PM's performance in the *Gutenberg Project* (see Table 22), as they both represent the industry-informed waterfall workflow with strict governance.

Commission in *General Collaborative Projects* in the crowdsourced environment was performed in a simpler way: the 'key' actors or the PM arranged tasks by saying '*Let us start revision*' or '*Part X is still unclaimed (or not revised), can someone do the work?*', and participants usually responded with direct action or claimed for the task in the chat board.

6.3.2.2.5 Social Interaction

Unlike crowdsourcing projects, in which social interaction (SO) appeared to be the most common dialogue act, the only *Social* interaction in the *Students' Team project* was apology (APM – 30 occurrences): e.g. apologise for the mistakes/problems in one's task or apologise for a delay. Other forms of social interaction were not used in the *Students' Team Project*, except for 'acknowledgement' such as '*thanks for your hard work and looking forward for our next collaboration*', which should also be counted as a form of social interaction. Moreover, because the communication in the *Students' Team*

Project was performed through email exchanges, it tended to lack opening greetings such as '*Good morning*', or concluding remarks such as '*Have a nice day'*, which are all social interactions if taking place in the chat board as dialogue segments. However, they were overlooked in favour of the other 'significant' topics in the email discourse, which is similar to the *Gutenberg Project*, another example of a formally-organised project in which the participants were not motivated by the need for social relatedness.

6.3.2.2.6 Maintenance

Acknowledgement (AK) is the second most frequent dialogue act in the *Students' Team Project* (see Figure 34), accounting for 57% of the interactions contributing to the *Maintenance* of the project. AK was used as a conventional expression in email exchanges to show gratitude towards one's work or help. It is a sign of politeness, which helps guarantee the friendly and effective collaboration between actors. AK was also present in the crowdsourced environment, but there it had more nuances, such as compliment, acceptance of information, and agreement on disputes.

'Confirm the receipt of file' (CF) was treated as an independent dialogue act in the *Students' Team Project*. As Matis (2014:169) indicated, it is the easiest way to guarantee that certain tasks were processed by the designated people at the right time. In the *Students' Team Project*, a reversed situation emerged: whenever a freelancer submitted a file or task, he/she would ask the PM to confirm the receipt of the file. On the one hand, this reminded the PM to check if there were any problems with the file in advance and allowed time for correction. On the other hand, in the final stage of the project, if something unexpected occurred – e.g. the PM claimed that a deliverable was missing, the freelancer could always use the 'CF' message to take a stand and then negotiate or re-submit the file. In the crowdsourcing projects, as all the tasks were performed through the online platform which was visible to all participants, CF was not a necessary act during the group communication.

Problem & Solution is another important subcategory in *Maintenance* interactions, which includes the following dimensions: problem report (PR – 79 occurrences), delay report (DLR – 16), solution to the problem (STP – 65),

assistance request (ATR – 16), and assistance arrangement (ATA – 7). The common problems were technology-related (i.e. challenges in using the designated computer-assisted tools), to which the PMs generally offered solutions through email, although sometimes on-site assistance was arranged, as well. In the crowdsourced environment, translation and technical problems were raised in the chat board, where usually the 'key' actors provided solutions.

6.3.2.3 Social Network Analysis

Just as in the crowdsourced projects, in the *Students' Team Project*, the whole team was seen as a network that contained individuals as clients, PMs, freelancers and instructors, and whose characteristics changed as the project progressed. Calculated by *weighted out-degree centrality*, the importance of individuals is suggested by the size of their nodes.

The *Students' Team Project* implemented a *waterfall workflow* in which the different stages of the project were clearly separated. Therefore, I visualised the network in accordance with the progress of the project. The aim was to explore the dynamic patterns of the interaction network and to investigate the contribution of actors. Figure 35 presents the graphs visualised from the social network analysis of the different project stages: setup – translation – revision – finalisation.

As mentioned in section 6.3.1, the *Students' Team Project* was divided into 10 teams. Therefore, in the visualised graph, we can see that, within the scale of the whole network, there are small networks representing each team centralised by the PM in charge, with surroundings consisting of freelancers. Freelancers are anonymised in accordance with their working language (e.g. ZH 1 refers to EN-ZH Translator 1, FR 1 refers to FR-EN Translator 1).

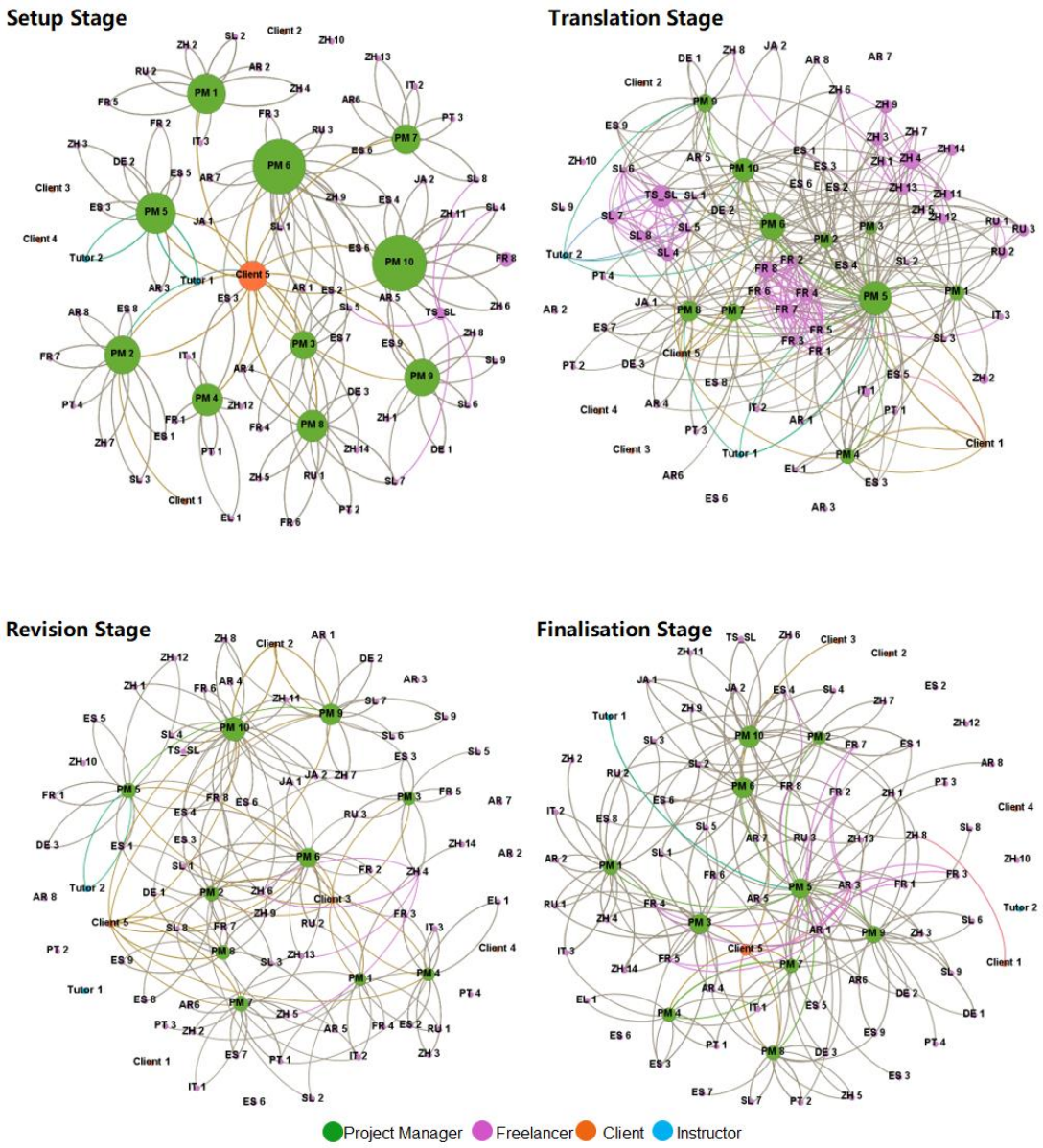

Figure 35. Visualisation of social network analysis of different stages in the *Students' Team Project*

The first implication drawn from my visualisations is the change of network scale. In the first project stage, we can see a relatively simple network in which the cluster of individual teams is clear, and the interactions are mainly performed between the PMs and the translators/subtitlers. In the translation/localisation stage, the network expanded, and the cluster of actors became more complicated due to the increased number of interactions. Several new clusters emerged within the network which were formed by the freelancers in accordance with their working language (i.e. the Slovenian group – SLs, the French group – FRs, the Chinese group – ZHs, and the

Russian group – RUs), who assembled to discuss how to deal with abbreviations, the translation of titles and names, and other terminology-related issues to ensure accuracy and consistency. Compared with the setup stage, the number of nodes around the PMs grew, which indicates that more participants were involved in the small networks of each team. At the same time, this growth of activity is also explained by the PMs who started approaching potential revisers in the translation stage to confirm their rates and availability for the task in advance.

The network dynamic then scaled down in the revision stage and reached a similar pattern as in the setup stage, but not identical because, as the project progressed, the nodes surrounding the PMs changed to include new teams of freelancers compared with prior stages, e.g. the revisers took charge of the project. In the finalisation stage, the cluster of actors in each team expanded again. More actors were involved in the small networks, during which translators accepted changes made by the revisers, finalised the tasks, and all freelancers submitted their deliverables.

The dynamic change of the visualised graphs is also reflected in the density of networks at different stages of the project (Table 30). Overall, the interactions and the participants' involvement in the translation (localisation) stage are more intense than in other stages of the *Students' Team Project*.

Table 30. Density of networks at different stages of the *Students' Team Project*

Networks	Setup	Translation	Revision	Finalisation
Density	0.084	0.166	0.067	0.073

Moreover, SNA reveals that interaction in the *Students' Team Project* mainly takes place in the forms of **one-to-many** (i.e. PM-to-freelancers) and **many-to-one** (freelancers-to-PM), which is caused by the hierarchical organisational structure and the closed working model. One notable exception occurred during the translation stage when a many-to-many pattern was needed for the discussion of translation-related issues between the freelancers in four language groups. The lack of more many-to-many

interactions is an interesting feature of the *Students' Team Project*, in which collaborative learning is encouraged but where students tend to seek the final answer from one individual instead of collectively deducing the answer.

The density of networks at different stages of the project further illustrated the interaction pattern: the closer the value is to 0, the sparser the network is, with no strong social relationship being established between freelancers. The network dynamics observed in the *Students' Team Project* resembles the traditional translation environment in which translators and other stakeholders are isolated from one another because of geographical and technological constraints. Information flow in such a condition is relatively private with low transparency, which also carries the risk of inefficient interaction. To be more specific, the 'one' actor – usually the PM – finds him/herself in a position of more responsibility, which can be challenging for beginners and trainees (as explained in the following paragraphs). For large multilingual localisation projects, such restricted interaction patterns weaken the cohesion between team members and reduce the coverage of knowledge exchanges. Because the *Students' Team Project* in this study was designed for learning purposes, a forum board was provided for freelancers to discuss translation problems. Many-to-many interactions thus occurred in the translation stage (although only within clusters of freelancers with the same working languages), but are less likely to be seen in the professional industry, making it obvious how, because of the absence of such complex interactions, professional localisation projects are seldom perceived as opportunities for learning.

Another characteristic in the visualised networks is the existence of 'key' actors, highlighting the importance of PMs for every stage of the project, with the occasional exception of certain freelancers working in certain languages who were more actively involved with the community to solve translation problems in the translation stage. Combined with the dialogue act analysis discussed in the previous section (6.3.2.2), the PMs' contribution covered planning, recruiting, coordinating, monitoring, problem-solving and ensuring the quality of the deliverables.

In the Language Services Industry and also in translation training programmes, the PM's contribution to multilingual localisation project is often overlooked and not many actors (e.g. freelancer or client) involved in the project are aware of the extended effort and time put in by PMs. To address this lack of knowledge, I used the student project time logs to assess the effort of the various actors involved in the project (Table 31). The result shows that the average time that the PM put into the project is **three times** more than any freelancer, however active these may be. For communication only, the PM's effort is more than **tenfold** the freelancer's. Consequently, I argue that the **PM is one of the key elements that influence the efficiency and the success of translation/localisation projects, and project management skill and effort needs to be valued more in translation training**.

Table 31. Comparison of time spent on project-related activities between the freelancers and PMs

Activities	**Average time spent (minutes)**	
	PM	**Freelancers**
Project planning	936	-
Communication with the client	95	99 (communication with the PM)
Communication with freelancers	1209	-
Translation	-	270
Subtitling	-	623
Revision	-	83
Financial	489	71
Troubleshooting	357	62
Other services (DTP, file managing, finalisation, quality check)	762	-
TOTAL TIME SPENT DURING THE PROJECT	**4132**	**1208**

Through the visualisation of the interaction network and dialogue act analysis, we are able to find out the trainees' strengths and weaknesses, as

well as examine the design and development of such project-based learning tasks, in order to improve translator training. For example, PM10 and PM6 appear to be the ‘key’ actors in the observed project (see Figure 35), but it is not always because the project success requires it. In fact, it is because PM10 experienced difficulties in recruitment, and PM 6 was assigned with extra tasks (image localisation). Moreover, PM10 self-declared that she was under great pressure (reflected in her email exchange with the instructor), which resulted in additional self-inflicted complications such as sending the files or messages to wrong addressees, or repeatedly and unnecessarily emphasising the project requirements to freelancers in email exchanges, which all led to extra time spent on the project.

To illustrate a different style of project management, my visualisations highlight that PM5 occupied a central position, as well, but for different, more productive reasons: he played a ‘coordinator’ role for the entire PM group – equivalent to a ‘senior PM’ role in the localisation industry. He was in charge of the weekly updates to the client, helped solve problems that other PMs encountered, updated all freelancers with the status of the whole project, and coordinated the overseas freelancers. Such forms of visualisation of team projects offer straightforward evidence for trainees to demonstrate and learn how an individual’s performance impacts the development of the entire project.

On-time delivery is often considered as the priority in localisation projects (Beninatto in Dunne, 2011:120). Through dialogue act analysis and project-log assessment, one can identify which project part is time-consuming and which procedures can be implemented to speed up the collaboration between freelancers and the LSP. In the case of the *Students’ Team Project*, repeated financial acts imply the trainee freelancers’ lack of proficiency in this area. To remedy this aspect in future projects, PMs may wish to consider providing a template to unify the structure of financial reports and using automatic calculation formulae to avoid mistakes, or using an online system which the freelancers cannot break. If such a problem is found in professional LSPs, one may need to introduce a project management system to reduce the excess time input.

Moreover, the PMs and freelancers in the *Students' Team Project* spent too much time communicating what was already clearly stated in the instruction file (e.g. dialogue acts PIC – project instruction clarify, DC – deadline clarify, and FFC – file format clarify in Figure 34). In future training, the instructor needs to increase the trainees' awareness of properly using the project brief/instruction file to avoid unnecessary time spent on such interactions.

At the same time, I found that trainees excel at safeguarding their benefits – for example, when negotiating rates (QTN), freelancers demanded the LSP to increase the rates offered to guarantee a profit for freelancers, and charged for any extra services they provided; in turn, the PMs also voluntarily proposed to raise extremely low rates for certain freelancers who were clearly undercharging, but also asked in return for the delivery of high quality output. By taking part in these mock projects, trainees learnt to establish professional manners for conducting themselves as businesspersons, which was reflected throughout their email discourse (e.g. dialogue acts AK – acknowledgement, CF – confirm the receipt of file, SR – status report in Figure 34). In addition, they also demonstrated their commitment and competence through collaborative problem solving (the request- and clarify-dialogue acts in *Information* category in Figure 34), which is in line with the crowdsourcing projects.

6.4 Summary

This chapter investigated one of the significant features of collaborative translation projects: the detail and organisation of online peer-to-peer interaction. It achieved this aim by first discussing the topic of conversational discourse using dialogue act analysis (DAA) and recognising its implications from several perspectives:

- the participants' intentions (i.e. what the participants wish to gain through specific dialogue acts, such as social – SO, or dialogue acts on requests, comment and clarifications).
- the influence on the progress of the project, mainly reflected through the dialogue acts with a stimulating and coordinating effect, such as social – SO, motivation – MT, commission – CM, revision request –

RR, status check – SC, and status report – SR.

- how dialogue acts can be used to monitor the project progress, predict problems, and make alternative plans to mitigate risk.

My annotation of the interaction data reflected the turn-taking feature through discourse units such as questioning, clarifying, discussing, and accepting. In addition to translation-related roles, participants in the crowdsourced environment also contributed in other positions, such as information provider, information requester, problem-solver, coordinator, and activator.

More importantly, this chapter provided evidence on the feedback and learning process found in the crowdsourced environment. It explained the giving and receiving of *elaborate feedback* in the form of translation problem solving, as well as *simple feedback* in the form of revision comment and compliment acknowledgement. The ongoing structure of turn-taking sequences is an indication of the degree to which information sharing and knowledge exchange influence the progress and quality of collaborative projects.

Since it is difficult to gain interaction data from the Language Services Industry, the *Students' Team Project* which was designed in the University of Leeds Centre for Translation Studies with significant input from Language Services Industry members working for *Sandberg Translation Partners Limited*, *Welocalise* and *Google* was taken as very similar in terms of communication activity to formal professional translation environments. Compared with the crowdsourced environment, communication was much more complex in the 'professional' localisation project: the result of the DAA in the *Student' Team Project* indicated that more dialogue acts were performed in the multilingual localisation (translation) project than in crowdsourcing projects. The extra interaction categories (*Recruitment* and *Financial* acts) are also frequently found in the professional industry, and it was interesting to notice that *financial* acts were the most frequent category throughout the *Students' Team Project*.

In the crowdsourced environment, social interaction (SO) and comment (CT)

were the top two most frequent dialogue acts, which reflects the nature of crowdsourcing where social-relatedness can be fulfilled. More importantly, SO is a characteristic of active participants, as it is natural for one to interact with others if they are working at the same time. Following SO, other dialogue acts can be performed smoothly, thus building a stimulating environment by involving more participants in the conversation. CT mainly contains the discussion of the project-related issues and translation problems and, given that one of the major reasons for the crowd to get involved in translation is the improvement of their own language and translation skills, CT has the same influence and effect as SO.

The previous chapter identified a network organisational structure in the four *General Collaborative Projects*, as well as a hierarchical organisational structure in the *Gutenberg Project (GP)*. My research found that GP was very similar in this respect to the *Students' Team Project*. The comparative study I conducted on these projects brought to light similarities in terms of the frequency of dialogue acts in the two projects sharing the same organisational structure (*GP* and the *Students' Team Project*), and showed that acknowledgement (AK) and status report (SR) are essential in the maintenance of formal projects. Moreover, both projects displayed an extensive amount of dialogue acts focusing on project instructions, which is explained by the implementation of strict project management, and especially by prioritising the quality of the deliverables.

Another finding of my comparative study was that social (SO) interaction was less common in the two projects with a hierarchical organisational structure, suggesting that both the participants in the *Students' Team Project* and the *Gutenberg Project* treated their work formally and seriously. Every dialogue thread in their interaction with peers had specific pragmatic functions, whereas the four *General Collaborative Projects* represent an informal and much more casual environment, in which the network organisational structure allows a much larger extent of autonomy to the contributors.

The dialogue acts regarding translation problems in the two environments share a similar effect in improving consistency and the output quality. The

difference is that translation problem-solving was centralised in the revision stage in the crowdsourced environment, while in the professional-like environment, translation problems were discussed in the translation stage. One possible reason is that, in the professional industry, having all problems solved during the translation phase reduces the effort and cost of revision. Combined with the result of the quality assessment in 6.4, my data highlighted the importance of implementing a longer and thorough revision stage in the crowdsourced environment, as it carried the majority of interactions regarding problem-solving and feedback, as well as the recognition and correction of translation errors.

Social network analysis (SNA) is an important part of this chapter, allowing the examination of the power relationship between participants, and the interaction pattern that captures the information flow. Through SNA, I first identified the 'key' actors in the interaction network of each project. For the crowdsourcing translation projects, the key actors are those who contributed the most when revising their peers' work (identified through workflow mapping in Chapter 5). This finding suggests that the 'active' participants in a crowdsourced environment are **more willing to** engage in a wide range of activities (in addition to their initially-assigned task) – **investing more effort and time in linguistic tasks** (i.e. translation and revision), and also **leading non-linguistic activities** (e.g. peer-to-peer interaction).

The key actors in the crowdsourcing network resemble the role of instructor (or tutor) in the formal training environment, as well as that of PM in the professional industry. Because they are actively involved in the production process, they are more sensitive to their peer's needs, and can therefore provide customised responses. The dialogue acts that key actors performed also made them the activators, information-providers, problem-solvers, and coordinators within the group, which optimised the collaborative model and enhanced the degree of knowledge exchange.

As opposed to the *General Collaborative Projects* in which the key actors were the primary participants in each project (e.g. an individual initially signed in as a translator), in the *Students' Team Project*, as well as the *Gutenberg*

Project, the PMs' contribution stood out. Following the analysis of the empirical data I had gathered, this emerged to be a characteristic of hierarchical organisational structures: the PM is assigned the heavy responsibility of controlling the progress of the project, which then means that **crowdsourcers requiring high quality and professional localisation projects need a very careful PM selection focused on a wide range of skills; moreover, the chosen PMs should be highly valued as they are one of the key factors that determine the successful completion of the project**.

Finally, the interaction networks visualised through SNA pointed out another difference between the crowdsourced environment and the professional-like environment: in the former one, a *denser* network was discovered, indicating a close and strong relationship between the participants. The interaction patterns included *one-to-many*, *many-to-one*, and *many-to-many* exchanges, which enhanced the collaboration between participants and expanded the coverage of knowledge exchange within the group. All of these are of great importance in the crowdsourced environment, as they allow participants to decide what information they want to receive, to designate the person who provides feedback, and to debate controversial points. By contrast, in the *Students' Team Project*, a *less dense* network was found because the network organised itself following a traditional hierarchical structure in which mostly *one-to-many* and *many-to-one* interactions were present.

Overall, my study showed that the interaction pattern in the crowdsourced environment works better for *self-organised* translation projects because it is much more suitable for exploiting the collective intelligence, collaborative learning, and collective problem-solving.

This chapter also demonstrated the feasibility of using DAA and SNA in Translation Studies, which brings new perspective to translation process research, particularly in the investigation of factors that influence the progress and quality of a translation/localisation project. These types of analysis can also help examine individuals' performance, as well as point out the strengths and weaknesses of participants, which will also be highly

valuable in a learning environment.

Chapter 7 Findings of the observation stage: 3 – translation quality assessment

Quality is one of the most frequently-questioned factors of crowdsourcing practice, but it tends to be viewed differently by practitioners as opposed to researchers. The former group has a tendency to evaluate quality with reference to the intended use of the translated product, while the latter community often speaks of translation quality in absolute terms, with little reference to the intended purpose and context of the translation. Examples such as localisation in *Mozilla Firefox* and *Facebook*, or translation in *Global Voices* and *the Rosetta Foundation* have already shown that crowdsourcing is suitable for language-related services, with 'fit for purpose' quality. However, this evaluation is mainly based on the translated output alone. Little research to date has investigated the production process of crowdsourcing using empirical data, which is a gap that this study has been designed to fill. More specifically, this study aims to address the discrepancy in evaluating the value of crowdsourcing translations between practical application and the academic world. In order to break the stereotypes still circulated regarding the quality of crowdsourced translation products, an error typology-based quality assessment was conducted to examine the translation product throughout the production process, and to explain how quality was controlled at several of the main production stages.

7.1 Data description – error annotation

Based on the simplified version of the *MeLLANGE* error typology, I annotated three versions of the translations produced in the observed crowdsourcing projects, mainly looking at the types and number of errors that were made in the first draft, and also at which errors were still left in the final product after two rounds of revision (self-revision and peer-revision). My findings were then used to compare the two current models used to organise projects in *Yeeyan*: the less-governed translation with a user-driven flexible workflow, and the strictly-governed translation with a waterfall workflow. The characteristics emerging from the analysis of the error annotation results

were also used to compare crowdsourcing practices across two translation genres: specialised translation and literary translation.

The error annotation was conducted using the *Translation Quality Assessment* (TQA) function of *SDL Trados Studio 2015 Professional*. Tables 32-36 illustrate the error distribution across translation versions which are displayed according to word count (around 1,000 words per part). For the four *General Collaborative Projects* belonging to the specialised translation genre, the activity ‘self-revision’ was not compulsory, and quality control largely depended on peer-revision. On the other hand, for the literary translation *Gutenberg Project*, due to the PM’s practical, hands-on and rather strict control, several rounds of peer-revision and self-revision were performed. The following presents the raw distribution of errors in each project, which will then be analysed in section 7.2. For the purpose of comparison across projects, the data were normalised on the basis of error count under each error type per thousand words (e.g. Part 1 represents the errors identified in the first thousand words of the source text).

Table 32. Error distribution in *Health Project I* (*HP I*)

HP I	**Content Errors**										**Language Errors**								**Style Errors**				**Hygiene Errors**			
	OM		AD		DI		TE		WTE		AM		AW		NN		RE		LT		TI		PU		CS	
	D	F	D	F	D	F	D	F	D	F	D	F	D	F	D	F	D	F	D	F	D	F	D	F	D	F
Part 1	2	1	0	0	2	1	7	0	0	0	0	0	2	1	7	2	0	0	0	0	1	1	3	3	0	0
Part 2	2	0	3	0	2	1	19	0	1	0	9	0	5	0	13	0	2	0	9	0	0	0	6	0	4	0
Part 3	0	0	2	0	12	0	23	0	0	0	12	0	4	0	22	0	0	0	8	0	0	0	1	0	1	0
Part 4	5	0	1	0	10	0	9	0	2	0	5	0	2	0	20	0	1	1	3	0	1	0	2	0	2	0
Part 5	2	0	1	0	12	1	13	3	1	0	10	5	3	1	10	1	0	0	3	2	1	1	7	4	3	2
Part 6	3	0	3	0	3	1	11	0	1	0	1	0	3	0	12	2	1	0	4	0	2	1	3	2	4	0
Part 7	2	0	1	0	7	0	10	0	0	0	3	1	0	0	7	0	3	0	0	0	1	0	2	0	1	0
Part 8	1	0	2	0	5	0	7	0	0	0	4	0	3	0	11	0	0	0	2	0	0	0	0	0	0	0
Part 9	0	0	0	0	1	0	7	2	1	0	2	1	0	0	5	0	1	0	1	0	0	0	0	0	1	0
Part 10	2	0	0	0	4	1	10	2	1	0	4	1	4	0	21	6	1	0	0	0	6	0	1	1	1	1
Part 11	1	0	0	0	1	0	8	0	0	0	0	0	4	0	10	0	0	0	2	0	2	0	0	0	0	0
Total	20	1	13	0	59	5	124	7	7	0	50	8	30	2	138	11	9	1	32	2	14	3	25	10	17	3

(D – Draft translation, F – Final output)

(OM – omission, AD – addition, DI – meaning distortion, TE – more suitable term available, WTE – wrong term, AM – ambiguous translation, AW – awkward syntax, NN – not idiomatic, non-native use of target language, RE – problem with register, LT – literal translation, TI – inconsistency within the text, PU – incorrect punctuation, CS – character spelling)

As indicated in Table 32, since no-one performed self-revision in *Health Project I*, the errors that were left in the final versions of each part are the result of a collective effort – peer-revision. The first notable finding is that NN represents the largest number of errors in the translation stage (25.65%), followed by TE (23.04%) and DI (10.97%). As for the performance of participants, T4's work (Part 3) contains the most amount of errors in the first draft, while T1 (Part 1) stands out to be the best translation with a total of 24 errors, mostly minor ones, such as TE and NN.

Table 33. Error distribution in *Health Project II* (*HP II*)

HP II	**Content Errors**															**Language Errors**												**Style Errors**						**Hygiene Errors**					
	OM			AD			DI			TE			WTE			AM			AW			NN			RE			LT			TI			PU			CS		
	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F
P1	7	7	1	0	0	0	1	1	0	8	8	4	0	0	0	1	1	0	0	0	0	4	4	1	0	0	0	0	0	0	3	3	0	1	1	0	0	0	0
P2	4	4	0	0	0	0	1	1	0	39	39	0	3	3	0	11	11	0	8	8	0	26	26	0	4	4	0	24	24	0	4	4	0	18	18	0	2	2	0
P3	3	3	0	2	2	0	4	4	0	20	20	0	3	3	0	8	8	0	6	6	0	12	12	0	2	2	0	1	1	0	0	0	0	10	10	0	0	0	0
P4	0	0	0	0	0	0	2	2	0	12	11	0	1	1	0	4	4	0	3	2	0	11	11	0	0	0	0	0	0	0	4	4	0	7	7	0	1	1	0
P5	2	1	0	0	0	0	0	0	0	17	11	0	3	2	0	4	4	0	2	2	0	16	13	0	0	0	0	2	2	0	7	7	0	5	5	3	0	0	0
P6	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	3	2	0	0	0	0	4	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Total	16	15	1	2	2	0	8	8	0	101	89	4	10	9	0	31	30	0	19	18	0	73	66	1	6	6	0	28	27	0	18	18	0	41	41	3	3	3	0

(P – Part, D – Draft translation, SR – Self-revision, F – Final output)

(OM – omission, AD – addition, DI – meaning distortion, TE – more suitable term available, WTE – wrong term, AM – ambiguous translation, AW – awkward syntax, NN – not idiomatic, non-native use of target language, RE – problem with register, LT – literal translation, TI – inconsistency within the text, PU – incorrect punctuation, CS – character spelling)

Table 33 shows the error distribution in *Health Project II*, where three participants spontaneously performed self-revision after translating. However, there is no significant reduction of errors after self-revision. The translators did not make any corrections in the hygiene category (PU and CS) during self-revision. TE, NN, PU are the top three errors in the draft translation (28.45%, 20.56%, and 11.55%, respectively). In terms of the participants' performance, T2's (Part 2 in the table) translation contains the largest

amount of error, while T1's (Part 1) translation has the fewest errors (a total of 25 errors), around a third of which are minor mistakes in TE.

Table 34. Error distribution in *Business Project I* (*BP I*)

BP I	**Content Errors**															**Language Errors**												**Style Errors**						**Hygiene Errors**					
	OM			AD			DI			TE			WTE			AM			AW			NN			RE			LT			TI			PU			CS		
	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F	D	SR	F
P1	2	2	0	0	0	0	3	2	0	32	23	1	0	0	0	5	3	0	6	5	0	22	14	0	1	0	0	6	4	0	10	7	0	7	6	0	2	1	0
P2	2	0	0	0	0	0	4	3	0	18	14	0	0	0	0	5	3	0	1	0	0	8	6	0	2	2	0	0	0	0	1	1	0	1	1	0	0	0	0
P3	6	4	0	1	1	0	5	5	0	24	18	1	1	1	0	6	6	0	3	2	0	13	12	0	1	1	0	3	2	0	3	3	0	4	3	0	0	0	0
P4	5	5	0	0	0	0	5	5	0	17	17	0	3	3	0	7	7	0	4	4	0	16	16	0	1	1	0	1	1	0	0	0	0	3	3	0	0	0	0
P5	2	1	0	2	1	0	5	4	0	46	38	0	1	0	0	7	5	0	4	3	0	16	13	0	2	1	0	0	0	0	5	2	0	7	6	1	1	1	0
P6	1	1	0	2	2	0	6	6	0	28	28	0	0	0	0	4	4	0	3	3	0	16	16	0	2	2	0	2	2	0	7	7	0	4	4	1	0	0	0
P7	0	0	0	0	0	0	0	0	1	2	2	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	1	0	0	0
Total	18	13	0	5	4	0	28	25	1	167	140	2	5	4	0	35	29	0	21	17	0	91	77	0	10	8	0	12	9	0	26	20	0	27	24	3	3	2	0

(P – Part, D – Draft translation, SR – Self-revision, F – Final output)

(OM – omission, AD – addition, DI – meaning distortion, TE – more suitable term available, WTE – wrong term, AM – ambiguous translation, AW – awkward syntax, NN – not idiomatic, non-native use of target language, RE – problem with register, LT – literal translation, TI – inconsistency within the text, PU – incorrect punctuation, CS – character spelling)

Table 34 illustrates the error distribution across the versions of *Business Project I*. TE and NN are the top two error categories (37.28% and 20.31%, respectively). Three translators performed self-revision in Part 1, 2, 3 and 5. However, they only managed to correct a few errors (in DI, TE, AW, NN, LT and TI).

Table 35. Error distribution in *Business Project II* (*BP II*)

	Content Errors															**Language Errors**												**Style Errors**						**Hygiene Errors**					
	OM			AD			DI			TE			WTE			AM			AW			NN			RE			LT			TI			PU			CS		
BP II	D	PR	F	D	PR	F	D	PR	F	D	PR	F	D	PR	F	D	PR	F	D	PR	F	D	PR	F	D	PR	F	D	PR	F	D	PR	F	D	PR	F	D	PR	F
	4	0	0	0	0	0	3	0	0	15	0	0	3	3	0	7	1	1	5	0	0	16	1	1	5	0	0	3	0	0	8	1	1	7	1	0	1	0	0

(D – Draft translation, PR – Peer-revision, F – Final output)

(OM – omission, AD – addition, DI – meaning distortion, TE – more suitable term available, WTE – wrong term, AM – ambiguous translation, AW – awkward syntax, NN – not idiomatic, non-native use of target language, RE – problem with register, LT – literal translation, TI – inconsistency within the text, PU – incorrect punctuation, CS – character spelling)

Table 35 shows the error distribution in *Business Project II*. Two rounds of revision were performed: peer-revision and third-party expert revision (invited by the PM). As the figures indicate, the translators cross-checked each other's work and managed to correct 90% of the errors. The expert reviser only corrected three wrong term errors (WTE) and one punctuation error (PU). Again, TE and

NN are the two major error types in the translation stage.

Table 36. Error distribution in the *Gutenberg Project* (*GP*)

	Content Errors										**Language Errors**								**Style Errors**				**Hygiene Errors**			
GP	OM		AD		DI		TE		WTE		AM		AW		NN		RE		LT		TI		PU		CS	
	D	F	D	F	D	F	D	F	D	F	D	F	D	F	D	F	D	F	D	F	D	F	D	F	D	F
Part 1	1	0	2	0	5	0	14	2	3	0	7	1	4	0	20	2	0	0	3	2	2	0	1	0	1	0
Part 2	1	0	0	0	9	1	10	2	6	0	9	0	2	0	21	1	0	0	3	0	0	0	2	0	1	0
Part 3	0	0	0	0	2	0	15	0	1	0	5	0	0	0	20	0	1	0	2	0	0	0	2	0	1	1
Part 4	1	0	0	0	1	0	13	1	1	0	4	0	2	0	18	0	2	0	1	0	1	0	6	0	3	0
Part 5	0	0	0	0	0	0	7	1	2	1	3	0	0	0	7	0	0	0	2	0	0	0	4	1	2	0
Part 6	0	0	0	0	2	1	8	0	6	0	4	0	1	0	14	1	0	0	2	0	0	0	11	2	1	0
Part 7	0	0	0	0	8	0	13	1	2	0	5	0	2	0	18	0	0	1	1	0	0	0	3	0	1	0
Part 8	2	0	0	0	2	0	17	1	2	0	3	0	4	0	18	0	0	0	3	0	0	0	5	0	3	0
Part 9	2	0	0	0	4	0	10	0	1	0	5	0	0	0	24	1	1	1	0	0	0	0	2	0	5	1
Part 10	1	1	0	0	3	0	6	1	1	0	4	0	1	0	23	0	2	0	1	0	0	0	2	0	3	0
Part 11	1	0	0	0	0	0	11	0	2	0	5	0	0	0	14	0	0	0	5	0	0	0	4	0	3	1
Part 12	0	0	0	0	1	0	8	0	1	0	5	0	4	0	16	0	0	0	1	0	0	0	8	0	3	0

Part 13	0	0	0	0	0	0	8	2	0	0	4	0	2	1	17	1	0	0	1	0	0	0	12	0	0	0
Part 14	0	0	0	0	0	0	14	0	4	0	0	0	0	0	17	0	0	0	1	0	0	0	3	0	2	1
Part 15	0	0	0	0	2	0	7	0	0	0	0	0	0	0	16	1	1	0	0	0	0	0	4	0	1	0
Part 16	0	0	0	0	1	0	9	0	1	0	1	0	1	0	13	0	1	0	2	0	0	0	6	1	2	0
Part 17	0	0	0	0	2	0	5	0	0	0	0	0	0	0	12	1	1	0	3	0	1	1	2	0	4	0
Part 18	2	0	0	0	0	0	4	0	1	0	1	0	3	0	16	0	2	0	1	0	0	0	3	1	3	0
Part 19	1	0	0	0	6	0	6	0	3	0	6	0	1	0	13	0	0	0	3	0	0	0	4	0	4	0
Part 20	0	0	0	0	3	0	8	0	3	0	6	0	3	0	21	0	3	0	2	0	0	0	14	2	1	0
Part 21	1	0	0	0	7	1	18	2	5	1	4	0	0	0	22	0	1	0	2	0	0	0	3	0	0	0
Part 22	1	0	0	0	4	0	13	0	3	1	4	0	5	0	25	0	1	0	4	0	0	0	5	0	3	0
Part 23	0	0	0	0	5	0	13	1	1	0	3	0	6	0	22	0	1	0	4	0	0	0	8	0	3	0
Part 24	1	0	0	0	3	0	6	0	1	0	3	0	2	0	25	1	2	0	3	0	1	0	5	0	2	0
Part 25	3	0	0	0	3	0	20	1	2	0	5	1	2	0	26	0	1	0	7	0	0	0	11	1	0	0
Part 26	2	0	0	0	0	0	12	1	2	0	2	0	2	0	19	1	2	0	2	0	0	0	9	1	3	0
Part 27	1	0	0	0	2	0	7	0	3	0	3	0	1	0	12	1	0	0	2	0	0	0	8	0	1	0
Part 28	0	0	0	0	3	0	11	1	0	0	4	0	0	0	17	0	2	0	0	0	0	0	6	1	1	0
Part 29	0	0	0	0	1	0	10	1	1	1	1	0	0	0	12	0	0	0	2	1	0	0	4	0	0	0
Total	20	1	2	0	79	3	303	18	58	4	106	2	48	1	518	11	24	2	63	3	5	1	157	10	57	4

(D – Draft translation, F – Final output)

(OM – omission, AD – addition, DI – meaning distortion, TE – more suitable term available, WTE – wrong term, AM – ambiguous translation, AW – awkward syntax, NN – not idiomatic, non-native use of target language, RE – problem with register, LT – literal translation, TI – inconsistency within the text, PU – incorrect punctuation, CS – character spelling)

Hospital Sketch is the title of the book whose translation project I observed in the *Gutenberg Project* (*GP*). My error annotation shows that language problems are more prevalent than content problems. Being able to translate a large literary text in an idiomatic and concise way is a major challenge for crowdsourcing participants (NN – 35.97%, AM – 7.36% in the first draft). TE (21.04%) and PU (10.90%) are the other two frequent error types. In *GP*, the revision procedures include peer-revision, self-revision and PM revision. The effectiveness of these procedures is analysed in the following sections.

The tables above merely describe the *raw* result of the error annotation in each project. The following sections present a more in-depth data analysis regarding the translation quality, the participants' competence, as well as the effectiveness of quality control methods.

7.2 Data Analysis

7.2.1 Holistic quality evaluation: comparison between specialised translation and literary translation in a crowdsourced environment

Prior to the holistic quality assessment, a source text analysis was conducted to assess its difficulty level based on two kinds of readability scores, the amount of specialised terminology, and other linguistic features (for details, please see section 3.2.4.1 in Methodology Chapter). The result was: for amateur translators in the crowdsourced environment, the translation of *GP* was estimated as a 'very difficult' task, three of the *General Collaborative Projects* (*HP I*, *HP II*, and *BP I*) were estimated as 'difficult' tasks, and *BP II* was seen as a 'normal' task.

Bar charts (Figures 36-40) below present a holistic evaluation of each project in terms of error distribution in its first draft and final published version which represents the raw translation competence of crowdsourcing participants and the effectiveness of revision procedures. As it shows, there is an extensive amount of errors in the draft version: ***Content*** and ***Language*** errors are the two major error categories in the four **specialised projects** (an average of 41% and 39%, respectively), while translation ***Language*** **errors exceed** ***Content*** **errors** in the **literary** project (48% and 32% of the total error counts in *GP*, respectively).

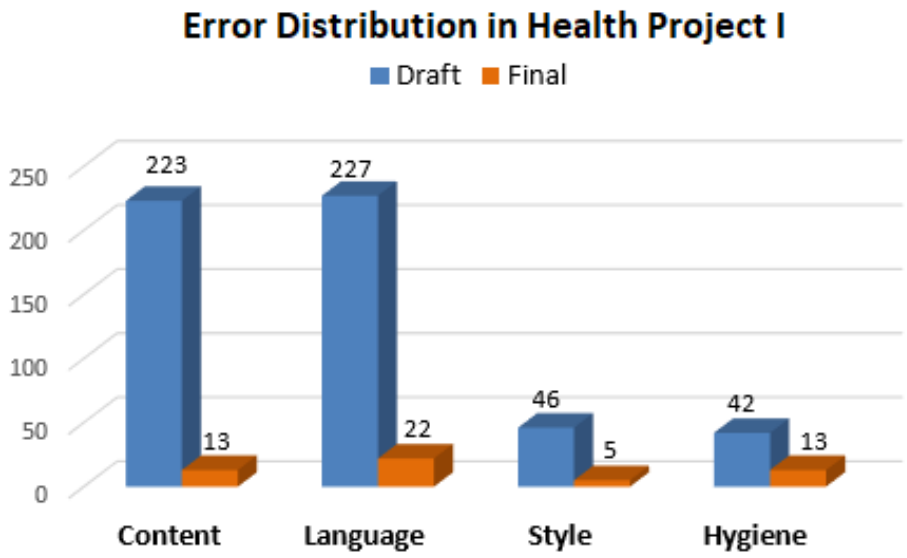

Figure 36. Error distribution across versions in *Health Project I* (*HP I*)

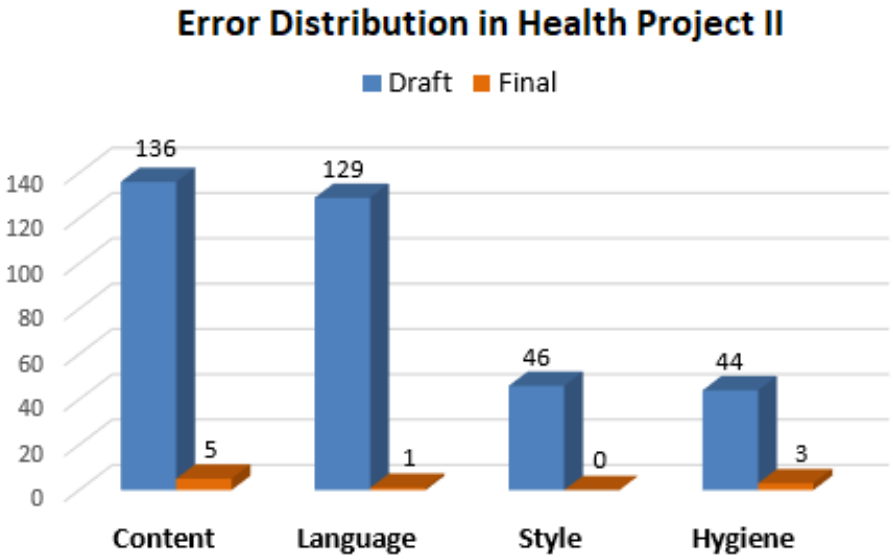

Figure 37. Error distribution across versions in *Health Project II* (*HP II*)

In *HP I*, 90% of the errors were corrected through peer-revision, while in *HP II*, 97% of the errors are corrected through both self-revision and peer-revision. Of the four error categories, *Hygiene* errors (incorrect punctuation and character spelling) had the lowest correction rate (69% in *HP I* and 93% in *HP II*). *Style* errors were fully revised in *HP II*, while around 10% of the *Style* errors in *HP I* remained uncorrected, which mainly includes

inconsistency of terminology within the text (TI). In terms of *Language* errors, 99% were corrected in *HP II* and 90% in *HP I* (most of the remaining errors were not idiomatic, non-native use of target language (NN) and ambiguous expression (AM)). The accuracy of the translation was improved through the reduction of *Content* errors (94% in *HP I* and 96% in *HP II*), and only minor terminology errors were left.

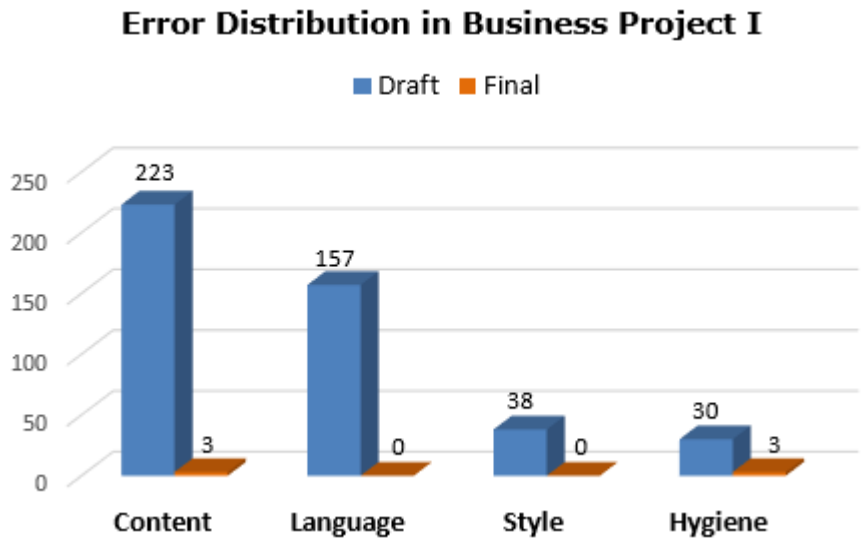

Figure 38. Error distribution across versions in *Business Project I* (*BP I*)

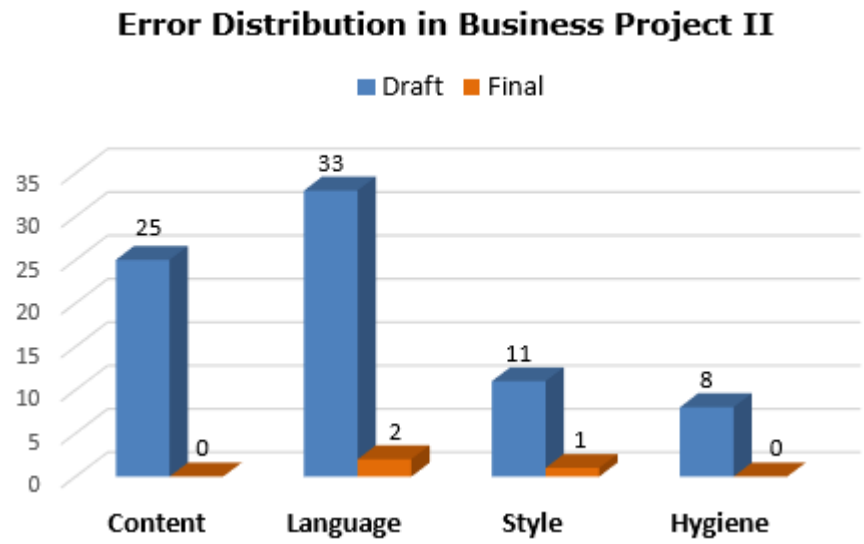

Figure 39. Error distribution across versions in *Business Project II* (*BP II*)

Compared with the two *Health Projects*, the *Business Projects* had higher correction rates in general (99% in *BP I* and 96% *BP II*). In *BP I*, all *Language* and *Style* errors were eliminated through both self-revision and peer-revision, and the correction rates of *Content* and *Hygiene* errors were 99% and 90%, respectively. The remaining errors in *BP I* were two minor errors of term choice (TE) and meaning distortion (DI) in one sentence, and three punctuation errors in *Hygiene*. In *BP II*, the accuracy of the translation was fully assured through the 100% correction rate of *Content* errors. There were no *Hygiene* errors left, either. The correction rates for *Language* and *Style* errors were 94% and 91%, respectively. The remaining errors were one ambiguous expression (AM), one not idiomatic, non-native use of target language (NN) and one inconsistency translation of a place name (TI).

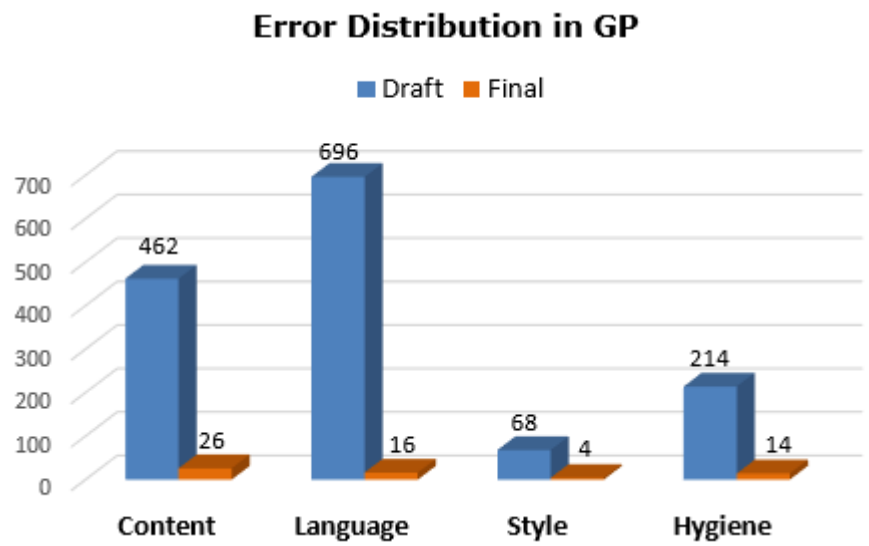

Figure 40. Error distribution across versions in *Gutenberg Project* (*GP*)

Although it was the most difficult task, *GP* still had a correction rate of 96%, in which 94% of *content* errors, 98% of *language* errors, 94% of *style* errors and 94% of *hygiene* errors were corrected, leaving no critical errors in the final translation.

Since this PhD project focuses on the crowdsourcing practices in China, it is important to examine whether the final product meets the industry criteria in the Chinese market. According to the *Specification for Translation Service* –

Part 1 (GB/T 19682-2005, GB/T19363.1-2008), the overall error rate of the final product shall not exceed 1.5‰. A formula is defined based on the difficulty level of the translation project, purpose of translation, category (weight) and count of the errors left in the translation (for details regarding the assessment, please see section 3.2.4.3, in the Methodology Chapter).

Table 37. Overall error rate in the final outputs across projects

Projects	*HP I*	*HP II*	*BP I*	*BP II*	*GP*
Overall Error Rate	1.5‰	0.49‰	0.23‰	0.6‰	1.7‰

Table 37 presents the overall error rate of each project, demonstrating that the final outputs of the four *General Collaborative Projects* have met the quality requirement of translation service in the Chinese professional industry. The *Gutenberg Project* was also very close to the quality standard in its specific category (i.e. translation for publication).

In general, the holistic evaluation indicates that the quality of crowdsourcing projects can be controlled through collaboration between participants. The figures testified that revision makes a significant improvement to the raw output produced in the translation stage. The high correction rates confirmed the effectiveness of collaboration and crowdsourcing participants' competence in solving translation-related problems. Although there are some minor errors left in the final version, given that these are translation projects that aim for knowledge-sharing, the final quality is sufficient for their purpose. On the other hand, in the case of *GP*, as a project for publication, there is still a little room for improvement.

Since the ultimate goal of this PhD project is to offer data-driven, evidence-based insights for future crowdsourcing practices, the challenge is to find out the weaknesses of the participants and the drawbacks of current workflows in order to come up with a solution that improves the overall competence of crowdsourcing participants and optimises the workflow of collaborative projects. The following sections examine the significant error types and the effectiveness of quality control procedures (preselection of contributors, self-

revision, and peer-revision).

7.2.2 Significant error types

To find out the weaknesses of participants (i.e. the particular knowledge and skills which non-professional translators lack), a simple statistical analysis was conducted to identify the common error types. Table 38 presents the mean and standard deviation values for each error type across translation versions in each project. As mentioned earlier, the data were normalised based on the error count under each error type per 1000 words for the purpose of comparison. It is important to note that in the case of *BP II*, it only shows the total count of each error type without mean and standard deviation because the word count of the source text is 971, for which normalisation does not apply. This, however, does not influence the examination of significant error categories.

The result shows that, in the specialised translation projects, among all the error categories, **TE** (*Content* error) and **NN** (*Language* error) are the top two error types with the most counts in the translation stage. AD (addition in the *Content* category) hardly occurs across all projects. In the final version, apart from **TE** and **NN**, **PU** (*Hygiene* error) stands out to be another significant one among the remaining errors. Surprisingly, TI occupies a very small portion of errors among all projects, which contradicts traditional criticism of crowdsourcing as being inconsistent.

Compared with the specialised projects, *GP* as a literary translation project shared some similarities: NN and TE are also the top two error types, and AD is the bottom one, while *Hygiene* errors (PU) also account for a much larger portion in *GP* in the first draft translation.

Table 38. Error type across versions in each project

Error Type	Mean (Standard Deviation)									
	HP I		***HP II***		***BP I***		***BP II* (total)**		***GP***	
	Draft	Final	Draft	Final	Draft	Final	Draft	Final	Draft	Final
OM	1.82 (1.40)	0.09 (0.30)	2.67 (2.66)	0.17 (0.41)	2.57 (2.15)	0 (0)	4	0	0.72 (0.84)	0.03 (0.19)
AD	1.18 (1.17)	0 (0)	0.33 (0.82)	0 (0)	0.71 (0.95)	0 (0)	0	0	0.07 (0.37)	0 (0)
DI	5.36 (4.25)	0.45 (0.52)	1.33 (1.51)	0 (0)	4 (2)	0.14 (0.38)	3	0	2.72 (2.46)	0.10 (0.31)
TE	11.27 (5.24)	0.64 (1.12)	16.83 (12.19)	0.67 (1.63)	23.86 (13.74)	0.29 (0.49)	15	0	10.45 (4.05)	0.62 (0.73)
WTE	0.64 (0.67)	0 (0)	1.67 (1.51)	0 (0)	0.71 (1.11)	0 (0)	3	0	2 (1.65)	0.14 (0.35)
AM	4.55 (4.11)	0.72 (1.49)	5.17 (3.66)	0 (0)	5 (2.08)	0 (0)	7	1	3.66 (2.14)	0.07 (0.26)
AW	2.73 (1.62)	0.18 (0.40)	3.17 (3.25)	0 (0)	3 (2)	0 (0)	5	0	1.66 (1.70)	0.03 (0.19)
NN	12.55 (5.92)	1 (1.84)	12.17 (8.26)	0.17 (0.41)	13 (7.1)	0 (0)	16	1	17.86 (4.66)	0.38 (0.56)
RE	0.81 (0.98)	0.09 (0.30)	1 (1.67)	0 (0)	1.43 (0.53)	0 (0)	5	0	0.83 (0.89)	0.07 (0.26)
LT	2.91 (3.08)	0.18 (0.60)	4.67 (9.50)	0 (0)	1.71 (2.21)	0 (0)	3	0	2.17 (1.54)	0.10 (0.41)
TI	1.27 (1.74)	0.27 (0.47)	3 (2.68)	0 (0)	3.71 (3.82)	0 (0)	8	1	0.17 (0.47)	0.03 (0.19)
PU	2.27 (2.37)	0.91 (1.45)	6.83 (6.62)	0.5 (1.22)	3.86 (2.48)	0.43 (0.53)	7	0	5.41 (3.41)	0.35 (0.62)
CS	1.55 (1.51)	0.27 (0.65)	0.5 (0.84)	0 (0)	0.43 (0.79)	0 (0)	1	0	1.97 (1.35)	0.14 (0.35)

Variance in the standard deviation (SD) value indicates the performance of participants. A higher SD suggests that the counts of a particular error type made by the participants are spread over a large range of values, e.g. in the draft translation of *HP I*, TE and NN are the commonly-made mistakes among participants (SD=5.24 and 5.92, respectively); a lower SD suggests that the counts of a particular error type made by the participants tend to be clustered closely around the mean, e.g. in the draft of *BP I*, RE and CS are not the common mistakes made by participants (SD=0.53 and 0.79, respectively). The drop of mean and SD between draft and final versions implies the effectiveness of the revision procedure in correcting a certain error type. For example, in the case of TE in *HP II*, the difference between the draft SD=12.19, and the final SD=1.63 signifies that the revision

procedure was effective in reducing the number of TE errors.

7.2.3 The effectiveness of the 'Preselection of Contributors'

Another quality control method implemented in the observed project (*GP*) is the 'preselection of contributors' through a sample test. It is also one of the features that differentiate the *Gutenberg Project* from the *General Collaborative Projects*. My quality assessment has already shown that both general participants and the preselected participants made a substantial number of errors in the draft translation, but what remains unclear is the actual translation competence of these two groups of translators, i.e. are the preselected translators better than the general participants? To find out, I conducted a T-Test by comparing the error count for different error types in the draft translation made by the translators in the *General Collaborative Projects* and the preselected translators in the *Gutenberg Project*.

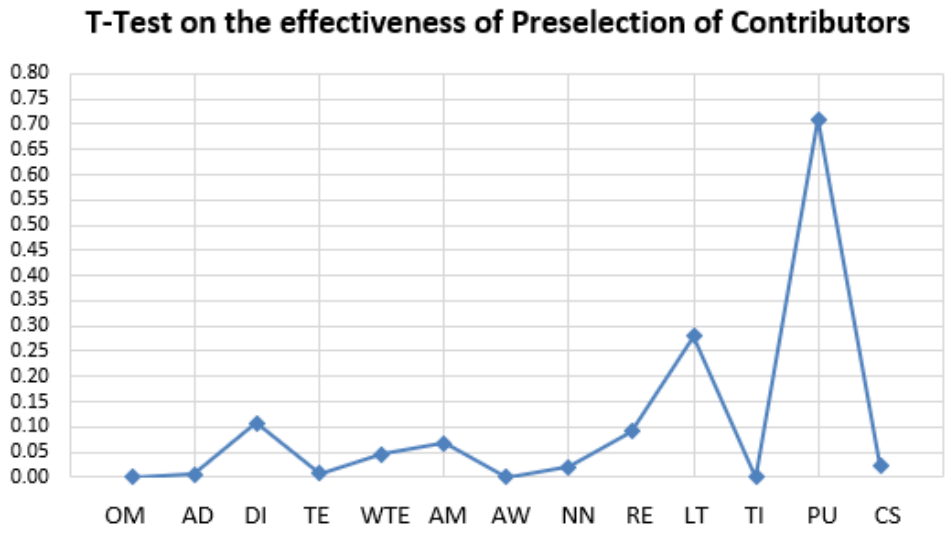

Figure 41. T-Test on the effectiveness of preselection of contributors

Figure 41 presents the result of the T-Test and demonstrates that the preselected participants do have higher translation competence than the general participants (who signed up for the projects freely). First of all, there is a significant difference in the error types of AW and NN (awkward syntax and not idiomatic, non-native translation, belonging to the category of *Language* errors), which implies that the tested translators master the target

language to a higher degree, an important advantage in a literary project which aims for publication and where accurately conveying the source content is as important as the ability to render the translation naturally in accordance with the rules and conventions of the target language.

Secondly, the tested translators are less likely to make errors of OM, AD, and TE, and WTE (belonging to the category of *Content* error), which is explained by the participants in *GP* being legally committed to the project, and acting according to professional ethics codes when performing the translation – i.e. do not omit the source information and do not introduce additional information. Another factor worth mentioning is that the count of OM and AD is much lower than other error types, which may result in bias in the T-Test. In terms of lexical errors (TE and WTE), as the literary text contains few specialised terms, the participants have fewer chances to make this particular type of error. Section 7.3.2.1 discusses the differences between the lexical errors in literary translation and specialised translation. Moreover, the preselected participants are also less likely to make errors in TI (inconsistency, belonging to *Style* errors) and CS (belonging to *Hygiene* errors) than the untested participants in the four general projects.

While the preselected translators in *GP* did not make as many *Content* errors, they made a similar number of DI (meaning distortion) errors compared with the participants in the *General Collaborative Projects*. Nevertheless, this problem was solved through peer-to-peer interaction due to the PM's proactive organisation of group discussions to solve the ambiguities in understanding the source text (discussed in Chapter 6).

The tested translators committed a similar amount of errors in RE (register), LT (literal translation) and PU (punctuation) compared with the non-tested participants. RE and PU are errors which can be easily ignored by amateurs if they did not receive formal translation training. One possible explanation for the common error – LT – may be that, for non-professional translators, when they come across a challenge, be it difficulty in ST understanding or TT rendering, it is natural to firstly produce a literal translation and revise it later.

Overall, the result of the T-Test demonstrates that a sample test is effective

in filtering out translators who have lower target language competence (reflected in the significant difference in the number of error types AW and NN made by the unselected crowd versus the selected GP translators).

7.2.4 Effectiveness of self-revision

The analysis in section 7.2.2 highlighted the significant error types that often occur in the crowdsourced environment, and the errors that can be corrected through revision procedures. The results of the questionnaire survey (Chapter 4) indicated that self-revision is one of the often-used quality control approaches in the *Yeeyan* community. In my experiment, eight participants in the observed projects were recorded performing self-revision (a *spontaneous* act by three translators in *HP II* and three translators in *BP I*, and a *compulsory* act by two translators in *GP*).

Figure 42 presents a bar chart of the percentage of total errors that were reduced through self-revision by the eight observed translators. In general, what is very important to notice is that self-revision does not prove as effective for correcting translation errors as peer-revision. This suggests that non-professional translators do not have the ability to spot their own mistakes in translation, but they are capable of correcting a useful number of errors made by other non-professional translators.

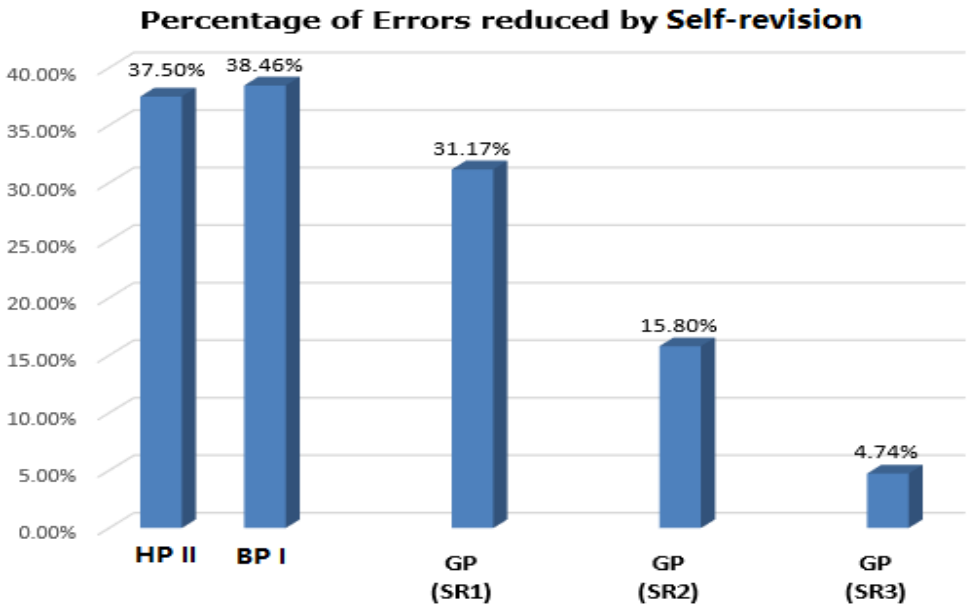

Figure 42. Percentage of errors reduced by self-revision across projects

A standard T-Test was further conducted to examine if there is any significant difference in translation quality after self-revision. Since edits for revision in the collaborative platform for the specialised projects were aggregated, self-revision usually occurred right after the translation stage. Therefore, I compared the draft translation and the version after self-revision. On the other hand, in the case of *GP*, self-revision took place after peer-revision to confirm whether to accept the corrections made by the other translator and to make other revisions, hence the self-revised translation was compared with the version prior to it (i.e. after cross-check and self-revision 2) (see the workflow mapping in Figure 24 in Chapter 5).

Figure 43 illustrates the result of the T-Test for the normalised parts with self-revision effort in *HP II* and *BP I*, which shows that there is no significant reduction of the error counts in nearly all error categories after self-revision, with the exception of TE and TI (inconsistency in *BP I*), where the p-value is less than 0.05.

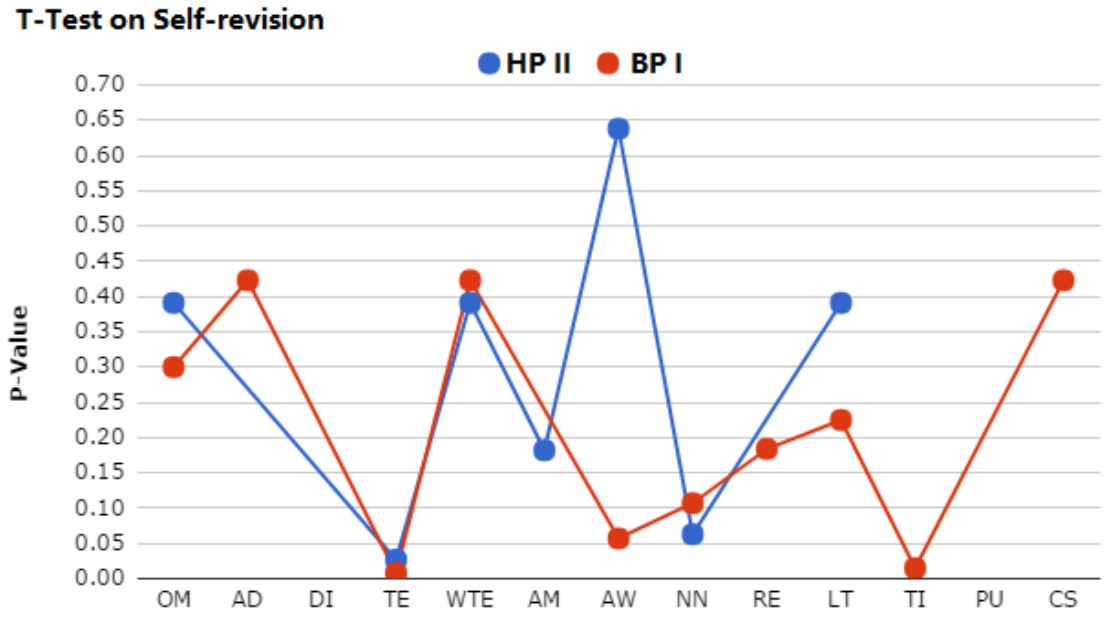

Figure 43. T- Test on self-revision in *Health Project II* (*HP II*) *and Business Project I* (*BP I*)

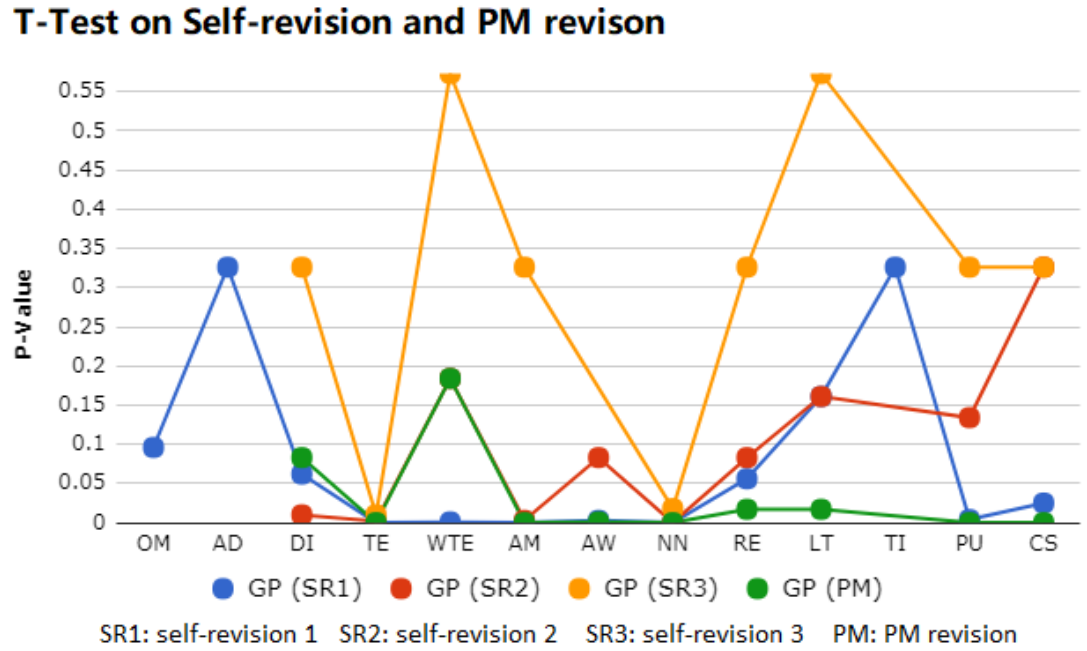

Figure 44. T-Test on self-revision and PM revision in the *Gutenberg Project* (*GP*)

For the *GP* (Figure 44), the error count in TE is significantly reduced between the three rounds of self-revision. The first round of self-revision (SR1) also

makes a significant difference on WTE, AM, AW, NN, PU and CS ($P<0.05$). As the project progresses, following the effectiveness of the cross-check process, fewer error counts occur as the control variables, which makes the influence of SR2 and SR3 less significant in the T-Test. Combined with the examination on the effectiveness of ‘preselection of contributors’ in section 7.2.3, I confirmed that the tested translators in *GP* have better target language competence which is reflected in the smaller AW and NN error count in the translation stage, and also in the effectiveness of self-revising AW and NN errors.

When evaluating translation quality in *GP*, I noticed that the PM’s final revision played an essential role in making positive corrections. Therefore, a T-Test was also conducted and the result shows that the PM’s final revision is much more effective than self-revision in reducing nearly all error types, particularly for *Language* errors (AM, AW, NN), *Style* errors (RE and LT), and *Hygiene* errors (PU and CS) (see Figure 44). The major impact made by the PM implies that, for a 'very difficult’ task as *GP*, the two translators’ performance is not sufficient for producing a high-quality translation even though they had passed the sample test in the recruitment stage. A thorough third-party reviser is still needed in this case.

The self-declared finding in the questionnaire survey demonstrated that general vocabulary, inconsistency within the text (TI), and misspelled character (CS) are the three most frequent error types that respondents found themselves correcting during self-revision; on the other hand, punctuation (PU), specialised terminology, and grammar are the least frequent error types (see section 4.5.3.1). This finding is also supported by the result of the T-Test on self-revision in the *General Collaborative Projects*, in which the participants only made significant corrections in TE and TI leaving PU as the least corrected category. Consequently, it appears that the respondents to my survey have a rather clear understanding of their self-revision competence. By contrast, the performance of the participants in the *Gutenberg Project* is much better. They made significant corrections in WTE, AM, AW, NN, PU and CS, possibly because they are preselected contributors who have better target language competence than the general

participants, which explains their effective corrections in the *Language* and *Hygiene* categories during self-revision.

7.2.5 Correlation between errors made and errors corrected

This section examines the competence of crowdsourcing participants from the perspective of peer-revision. My assumption was that, although participants make errors in their draft translation, they have the ability to correct a similar amount of errors in terms of error types made by other members of the crowd. A correlation coefficient test was therefore applied to compare the errors that one makes with the errors that one corrects within an identical word count range.

To address the bias in the analysis – for example, fewer corrections are made in a part where fewer errors exist – I compared the *percentage* of errors that were made under each type per thousand words with the *percentage* of errors were corrected under each type per thousand words.

The result is shown in Figure 45 and Figure 46. Nodes with large residuals (fitted deviation) are presented in different colours (the ‘green’ nodes refer to the error type NN, ‘blue’ nodes refer to TE, and ‘yellow’ refers to OM). The residual is the distance between the nodes and the fitted line, indicating the participants’ varied ability in correcting different types of errors.

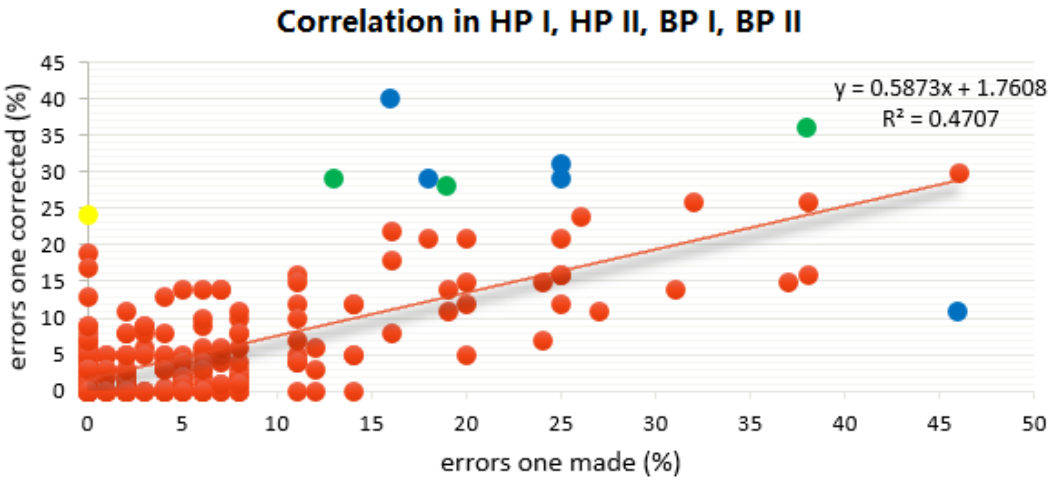

Figure 45. Correlation between the errors made and errors corrected in the *General Collaborative Projects*

A positive relationship is found in each project. For the four *General Collaborative Projects* (Figure 45), the Pearson correlation coefficient is 0.69, while for the *Gutenberg Project* (Figure 46), the correlation coefficient is 0.45: the result suggests a stronger correlation in the four *General Collaborative Projects* than in the *Gutenberg Project*. Although there is an entrance examination in the latter one, it does not demonstrate that the 'filtered' (preselected) participants are better than the general participants in performing the **peer-revision** task.

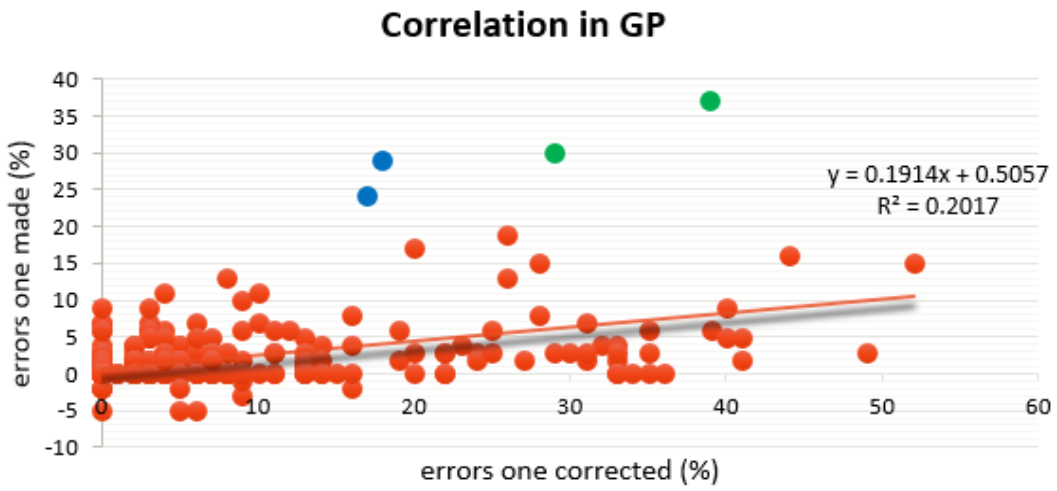

Figure 46. Correlation between the errors made and errors corrected in the *Gutenberg Project* (*GP*)

As can be seen in the Figure 46, there is a small portion (5%) of revisions made by the translators in *GP* which are negative because the two translators introduced errors during peer-revision due to their disagreement in understanding the source text. A possible explanation can be that the source text is a literary text published in 1863 which tells the story during the American Civil War, contains expressions in Latin, regional dialect, and many complex sentences full of rhetorical phrases of the time. As indicated by the translators in the follow-up interview, due to the style and characteristics of the source text, they had difficulty in understanding the content accurately. Hence, their effort in peer-revision seems to be not as effective as the participants' performance in the *General Collaborative Projects*.

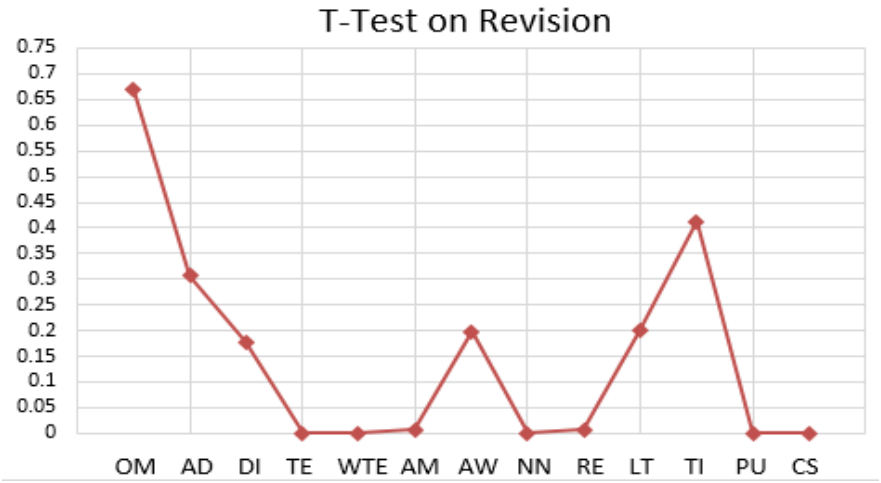

Figure 47. T-Test on the effectiveness of peer-revision across error types in all projects

A T-Test was further conducted to see which are the error types that the participants are good at spotting and correcting. As shown in Figure 47, of all the projects, the participants are particularly good at reducing TE and WTE (*Content* errors), as well as AM, NN and RE (*Language* errors). The p-values of TE and NN are the lowest of all, leading to the conclusion that TE and NN are recognised as the two most significant error types during peer-revision.

In general, the correlation analysis verified the crowdsourcing participants' competence in peer-revision. They make different kinds of errors during the translation stage, but they are also able to spot and correct these kinds of errors when they appear in others' work.

7.3 Discussion

7.3.1 User-focused approach: enhanced collaboration

Desjardins (2017:24) discussed online collaborative translation through the relationship between online social media and translation. She pointed out that current research in Translation Studies surrounding 'how platforms [such as *Facebook* and *Twitter*] are translated might not generate novel insights' and instead she suggested focusing on the 'crowd' that produces the translation work. To be more specific, she recommended profiling a sample

population of the ‘translator crowd’ to see what type of competencies these individuals possess (ibid:24), which is one approach I implemented in my project.

A common feature in crowdsourced practices in the West is that volunteers are usually asked to merely perform the ‘translation’ activity, and their participation ends when the translation is finished (e.g. *Global Voices*, *Mozilla Firefox*). Afterwards, it takes time for revisers or supervisors to pick out and amend the poor output, which is an essential criticism for crowdsourcing in general (Rajala et al., 2013). However, my research highlights that the situation may change if the crowd are empowered to collaborate more fully – by allowing participants not only to perform the translation task and/or vote on others’ translations, but also actually spot errors and correct others’ work.

The results of my error annotation process demonstrate that:

1. All participants in the crowdsourced environment are capable of producing draft-level quality, maintaining commitment to the project and very rarely abandoning their assigned tasks;
2. Translation quality can be significantly improved through revision, particularly via peer-revision.

The crowdsourcing model in *Yeeyan* has enhanced collaboration between participants to the largest extent. As discussed in Chapter 4, one of the strong motivations that drive people to take part in crowdsourcing translation is the opportunity to practise and improve their language and translation skills. The collaboration mode of having people translate, allowing people to view and edit others’ translations, and encouraging peer-revision at all times is a sensible and practical approach that helps participants maintain this motivation. My research has empirically proved the effectiveness of implementing a *high* level of accessibility of peer contributions in the crowdsourced environment. Participants are exposed to their weaknesses through the feedback provided by peers. Regardless of whether this feedback is presented in the form of revision comments or revision edits, it still corresponds to a mentoring and learning process through collaboration.

Moreover, **for very difficult texts that require a higher quality standard** such as the *GP*, **an expert reviser is necessary**. The expert revision can be achieved through the involvement of the project manager (PM) – provided the PM is a highly-experienced translator within the community – or the introduction of an expert reviser. The result of my data analysis suggests that more involvement is better in the production process. In other words, a **high accessibility of peer contributions** is useful in assuring translation quality by allowing participants to view, discuss and revise peers' work.

7.3.2 Exploring the crowd's weaknesses

Participants in the crowdsourced environment are expected to produce translations in accordance with the purpose of the project. The analysis of the error annotation stage of my project indicates that TE (more appropriate term available) and NN (not idiomatic, non-native use of target language) are the top two error types made in the translation stage, and are also the ones being corrected most often in the revision stage. TE and NN belong to the two broad error categories – *Content* and *Language*: 'content' refers to the accuracy of the translation, while 'language' caters for the fluency of the translation. Both parameters are essential in achieving the equivalence effect for the target readers.

Since TE and NN are the most commonly-made mistakes among amateurs in the crowdsourced environment, it is necessary to further investigate why these errors occur – i.e. the crowd's weaknesses in translation. On the one hand, it helps structure these errors with detailed elaboration; on the other hand, it may suggest a path for future practices by pointing out what needs to be paid attention to when using a crowdsourcing product, especially under the common circumstance that the 'crowd' are only recruited for the 'translation' task. Most importantly, it can offer insights into what kind of training or guidelines should be provided to improve the crowd's competence.

7.3.2.1 Subcategories of TE

The category of TE (more appropriate term available) refers to all lexical problems, including both specialised terminology and general terms. When annotating errors, I found that lexical problems are present in various ways. They are summarised into 10 subcategories specifically for EN-ZH translation: (1) ambiguous meaning, (2) collocation, (3) conjunction, (4) contextual adaptability, (5) frequency, (6) measure word, (7) pronoun, (8) lexical property, (9) repetitive and not concise, and (10) specialised terminology. Below is a bar chart that illustrates the percentage of TE types across text types in the translation stage.

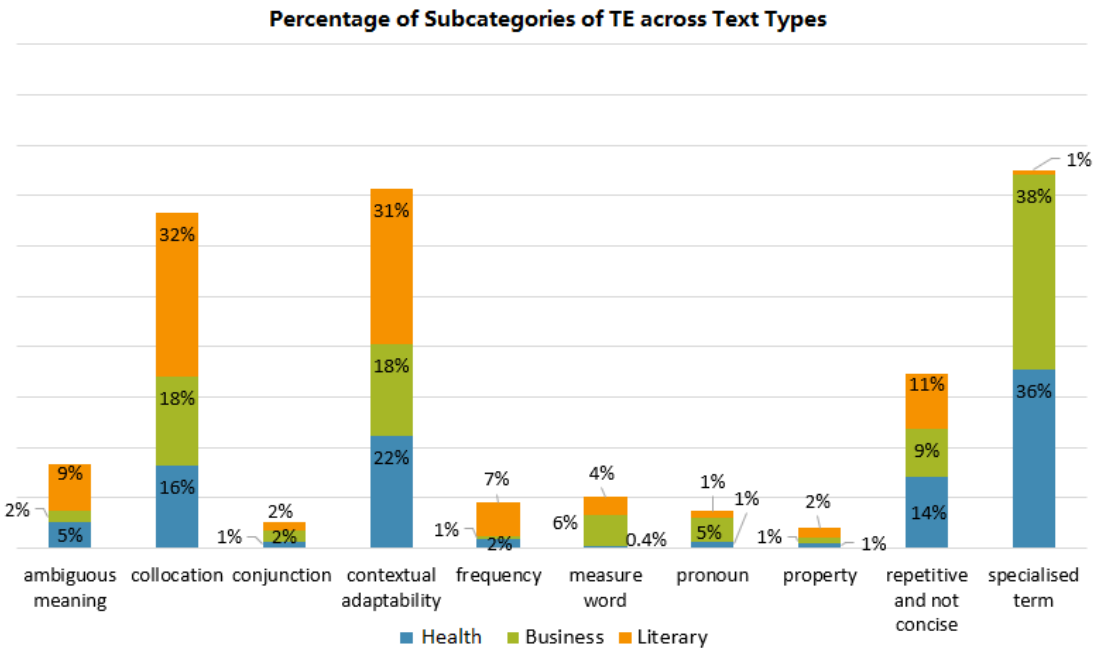

Figure 48. Subcategories of TE across text types

As Figure 48 indicates, the percentage of subcategories in TE varies in different text genres. For specialised projects (Health – *HP I* and *HP II*, and Business – *BP I* and *BP II*), nearly half the lexical errors are related to **specialised terminology** (38% in Health, and 36% in Business), while in the literary project (*GP*), few specialised terminology errors are made (1%). As mentioned in section 7.2.2, a shared glossary list proved to be an effective approach for reducing specialised terminology errors across all the observed

projects. The participants first added the bilingual specialised terminology they encountered to the shared glossary list during or after the translation stage, and then when they switched to the role of reviser, they consulted the glossary list that ensured both accuracy and consistency.

Collocation is the second most frequent TE type of all TE errors among all text types (32% in Literary, 16% in Health, and 18% in Business). Understandably, having the right collocation contributes to the fluency of the translation, and more specifically to the native use of the target language. This is even more important when working on literary texts, given the aim and nature of their translation. Below are some random examples of collocation errors from the observed projects (the largest available Chinese corpus in *Sketch Engine* – zhTenTen 2011 – was used to compare the frequency of collocations). Table 39 presents examples of collocation problems randomly selected from the literary project (*GP*), demonstrating the added value of revision: **the collocations in the final translation are more frequently-used than their initial versions proposed in the draft translation**. Consequently, for community-based crowdsourcing initiatives, the crowdsourcer should consider providing training or guidelines on the use of corpus tools to inform strategies for solving lexical errors (collocations in particular) and to help improve the crowd’s translation competence.

Table 39. Examples of collocation errors in TE

Source Term	**TE in Draft translation**	**TE in Final**	**Frequency score in Corpus** (Draft: Final)
fulfilment of my own dark prophecy	满足了我的黑暗预言	印证了我的黑暗预言	0: 5.06
a realizing sense of the fact	清晰的体认	清晰的认识	0: 6.03
returned to the path of duty	坚守着职责的道路	坚守着职责的本分	0: 0.36
eyes spiritless and sad	眼光无神而忧伤	眼睛无神而忧伤	0: 1.95
strong, properly trained, and cheerful men	强壮快活、训练有素	强壮乐观、训练有素	0: 3.51
patriotic fancy	爱国想象	爱国情怀	0: 3.39
bedizened with a bit of finery	配饰些华丽的装扮	佩戴些华丽的装扮	0: 7.26

The third frequent TE type is **contextual adaptation** (31% in Literary, 22% in Health and 18% in Business), which refers to two considerations when making a word choice: 1. faithfulness – the selected term should deliver the right meaning in accordance with the context of the source text; and 2. elegance – the selected term should suit the text genre. For instance, in a literary text the term may need to be elegant, while in a specialised text, the same term would need to be precise and formal. The two considerations are derived from the translation standard proposed by the Chinese scholar Yan Fu – Xin, Da, Ya (信, 达, 雅 – faithfulness, expressiveness and elegance) (Munday, 2012:49). Figure 48 indicates that, out of all TE errors, compared with the specialised texts, contextual adaptability occurred much more frequently in the literary text. As declared in the follow-up interviews, two translators in the literary project (*GP*) did not receive formal translation training (one majored in biology and the other in law), and therefore may not have had the awareness of taking all these considerations into account when making appropriate word choices.

Repetitive and not concise is the next frequent TE type, which refers to the situation when a source term is repeatedly translated using different forms, or the translation reads as verbose and complicated. Like collocation and contextual adaptation, it occurred more frequently in the literary text than in the specialised texts. It is rather understandable that the amateurs make this particular type of error because their main focus when translating is to convey the source text meaning fully in more of a word-for-word than a sense-for-sense manner. Moreover, the crowdsourced translators do not appear to master the necessary research and evaluation strategies to correct such errors while translating or during self-revision. Nevertheless, it is the kind of problem that can be easily spotted in peer-revision, as indicated by its numerous mentions in the revision comment (RC) posted in the chat board.

Conjunctions are important function words that determine the fluency of a text. Conjunctions in Chinese can be divided into two major types: those coupling words or phrases and those linking clauses and sentences (Yip and Don, 2004). Conjunction errors usually result in a non-native and unnatural use of the target language (NN). In my research I noticed that conjunction errors make up 5% of all TE errors (1% in Health, 2% in Business, 2% in Literary), and are explained by the tendency of participants to focus too much in the translation stage on the literal meaning of the source text, consequently ignoring or forgetting to add conjunctions that will improve the fluency of the translation. This problem can be addressed through 'translation guidelines': for instance, the crowdsourcer can compile a set of guidelines on common errors such as 'conjunction' and disseminate it within the community for the purpose of raising the crowd's awareness and knowledge of using conjunctions while translating. Since the respondents to the questionnaire survey have already demonstrated that their primary motivation is to improve language and translation ability, they will be willing to receive such a resource and follow its instructions. When initiating a crowdsourcing translation project, it is important for the PM to provide guidelines emphasising common problems (e.g. instructions on conjunctions, punctuation rules, requirements for specialised terms) to avoid unnecessary translation errors and reduce the cost and effort during revision.

Measure word, **pronoun** and **lexical property** are three Chinese-specific lexical problems. Each of them accounts for less than 10% of TE errors in all text types. One of the most distinctive features of a Chinese common noun is that a measure word is normally used in conjunction with a number or demonstrative. In some cases, the measure is a quantifier (e.g. *a* mountain, *two* flowers), and in others it is a universal or standard measure, which is generally associated with material nouns (e.g. *a drop* of water, *a cup* of tea). In Chinese, material nouns are common nouns which have their own specific measure words. Translators need to have the awareness of transferring the English quantifier and measure word to a Chinese-customised measure word in order to achieve naturalness in the target language (Yip and Don, 2004). This error appeared more frequently in the business and literary text than in the health text (0.4% in Health, 6% in Business and 4% in Literary).

Pronoun error mainly refers to personal pronouns. In Chinese, there are three types of personal pronouns: first person (I – 我, we – 我们), second person (you – 你, you – 你们 and the polite form of you – 您), and third person (he – 他, she – 她, it – 它, they – 他们/它们) (ibid). The error in this study lies in the misuse of third person pronouns. '它' [it] refers to a non-human object, e.g. animals, material nouns, and abstract nouns. Participants often misuse '它' [it] instead of '他' [he] and '她' [she], which may because they have not developed a habit for using them accurately in a formal context. Another possible reason is that the pronunciation of the three third person pronouns is the same. If translators pay insufficient attention to checking their keyboard output (when using the pronunciation-based Chinese input method '*pinyin*'), this will lead to the unconscious occurrence of such errors.

Lexical property, specifically for adjectives, can be divided into three groups: commendatory, neutral and derogatory term. In general, terms are usually affiliated with sentiments, making them positive, negative or neutral while having the same semantic characteristics. Translators need to choose a term with suitable properties so that it successfully delivers the source meaning and the author's attitude towards an entity. This error occurred more

frequently in the Literary text (2%) than in the specialised texts (1% in Health and 1% in Business), possibly because it is somewhat associated with the nature of the source text – in general, there are more adjectives in a literary text, and hence a higher chance of making the mistake.

Ambiguous meaning refers to when the selected term cannot fully covey the meaning and intention of the source term and sometimes can cause ambiguity in the readers' interpretations. The occurrence of this error is less than 10% in all text types.

Frequency in TE signifies that the selected term is less frequently used considering the source context and text type. When translating a text for knowledge sharing, the participants should avoid using difficult, complex and less-frequent words.

7.3.2.2 Subcategories of NN

Among all NN (not idiomatic, non-native use of target language) errors, subcategories include (1) word order, (2) redundant expression, (3) complex sentence structure, (4) missing or misused function word, and (5) other miscellaneous problems that influence the smoothness of translation. NN can also be caused by literal translation (LT) and awkward syntax (AW).

Figure 49 is a bar chart that shows the percentage of NN types in different text genres. **Missing or misused function words** represents the most frequent problem that makes the translation sound unnatural and non-native (58% in Business, 41% in Health and 37% in Literature). Function words refer to Chinese-specific particles and conjunctions, and their absence does not make the syntax or grammar of the translation wrong, but they are still highly influential in the audiences' reading experience. Native-Chinese speakers are accustomed to the existence of function words, and if the translation simply conveyed the meaning of the source sentence but failed to add the function word to bridge the language differences between English and Chinese, the fluency of the translation is significantly diminished in the eyes of the target audience.

For example, because there is no conventional way of expressing

grammatical tense in Chinese as opposed to English, the particle *le* '了' is suffixed to a verb to emphasise a completed action. The particle *zhe* '着' is suffixed to a verb to show a progressive action or continuous state. By adding these function words, readers can understand the status of a verb/action in a Chinese-customised approach. Other function words include conjunctions between clauses, modal particles, auxiliary verbs, adverbs, and prepositions, which all have their unique meaning in Chinese. Interestingly, T2 in *GP* stated in the follow-up interview, that after translating for a long time, he had found himself having difficulties in using Chinese properly. He concentrated too much on the source language (English) and forgot the conventions in Chinese, such as the use of function words.

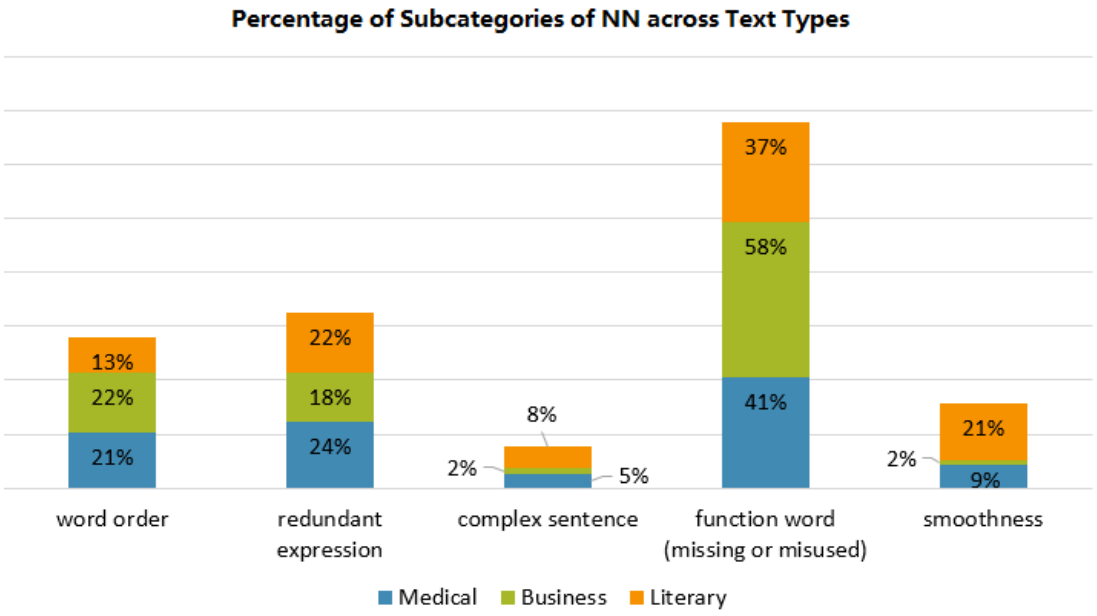

Figure 49. Subcategories of NN across text types

Redundant expression is similar to the 'repetitive, not concise' error in TE: it signifies that the translation of a clause or sentence reads as too verbose and there exist simpler and clearer ways of expression. **Complex sentence structure** is another problem caused by the difference between Chinese and English, which can lead to ambiguity in translation. Translators need to establish the awareness of breaking an English complex sentence into several sentences in Chinese by using conjunctions or suitable punctuation

to achieve ‘expressiveness’ in the translation. It is a challenge often recognised by the participants in *GP* in the follow-up interviews.

Word order is the third frequent NN type, and it is most likely to be mistaken by amateur and beginner translators. Word-for-word translation is the first step in acquiring translation skills and is often observed in the crowdsourced environment: as novices translate, their natural instinct is to follow the word order of the source text, even though it may not be suitable in Chinese. Having said that, it is important to notice that the correction of word order problems is the easiest task to perform, and all errors related to word order were identified and revised by participants.

There are other non-idiomatic and non-native uses of the target language that cannot be categorised in the above types, and were therefore combined in the category ‘**smoothness**’ in Figure 49. The literary text has the most amount (21%) of miscellaneous causes for NN due to the complex sentence structure in the source text, and the PM’s final revision in *GP* was extremely valuable in correcting these errors (see Figure 44).

7.4 Summary

This chapter examined crowdsourcing projects under different settings, using empirical TQA as a methodology – one which crowdsourcing research in the Translation Studies domain rarely uses due to its time-consuming nature, as well as the general absence of first drafts, revised translations and final accepted translations to compare and contrast. By researching a collaborative platform which stores all of these versions, I was able to produce a novel study which investigates the issue of quality and which proposes optimal organisational model for projects with different aims.

The error annotation of the draft translation demonstrated that the **‘raw’ output of the crowd needs vigorous revision to ensure the quality of the final product**. **In that respect, peer-revision is** so far **the most efficient approach to improving translation quality**. **Participants** may not perform well in spotting and correcting their own mistakes in the self-revision stage, but they **are capable of finding others’ errors and make suitable**

revisions. My research verified the assumption that the degree of collaboration influences the quality of the output – more involvement is better than less in the translation process – and that can be achieved through the process of crowdsourcing – i.e. enable *high* accessibility of peer contributions so that collaboration is enhanced. Amateurs can produce adequate solutions to translation problems through collaboration in the form of group discussion and aggregated revision edits.

Moreover, this study proved that the 'preselection of contributors' is an effective approach for quality control, as well. A sample test is useful in selecting contributors with higher target language competence, and my statistical analysis demonstrated that the tested translators in the *Gutenberg Project* made fewer errors in AW and NN compared with the untested participants in the *General Collaborative Projects*.

A comparison of error types across versions indicated that the translation of both specialised and literary texts is feasible in the crowdsourced environment. The differences in the significant error types pointed out the weaknesses that amateurs possess: terminology is the most evident obstacle for specialised texts, but my research showed that having a shared glossary list is effective for reducing inconsistency and in improving content accuracy. Participants who performed the revision activity tended to respect the glossary list.

Language errors (NN) are also very frequent across all projects, but even more so in the literary project. The result of the error annotation across versions in *GP* demonstrated that the PM corrected a significant number of language errors, highlighting the importance of this role, and also demonstrating the necessity of a third-party reviser for improving the fluency of the translation.

Last, but not least, the most significant contribution of TQA in these observed projects is the re-identified subcategories of significant error types in TE and NN. They highlighted specific obstacles which future organisers and amateurs need to pay attention to in the crowdsourced environment. More importantly, they represent a detailed EN-ZH specific TQA model for

translator training – to help trainees identify flaws in their own translation during self-revision, and to instruct trainees on how to revise others' work with structured feedback.

In the next chapter, I combine the findings from the questionnaire survey, communication analysis, workflow mapping, and quality assessment to present a more thorough picture of the research outcomes and implications.

Chapter 8 Conclusions

This chapter summarises the findings of the investigation into the initial research questions, makes suggestions for the future study and management of crowdsourcing initiatives, as well as acknowledging the limitations of the current research and suggesting directions for future work.

My research focused on crowdsourcing translation in the context of the *Yeeyan* community. In light of Zhao and Zhu's (2012) theoretical framework of crowdsourcing research, it explored the 'Who', 'Why' and 'How' dimensions of this particular crowdsourcing initiative, and it improved that model by adding a further dimension which is extremely relevant in the context of the Language Services Industry: the 'Quality' of the crowdsourcing outputs.

I designed a questionnaire survey for the members of the *Yeeyan* community to investigate the 'Who' and 'Why' dimensions using self-declared data. I also observed the two major types of collaborative translation projects in *Yeeyan*: a *General Collaborative Project* that focuses on popular news article translation for knowledge-sharing; and the *Gutenberg Project* that focuses on book translation for publication. The two types of projects represent different text genres: the former one represents specialised text with informative features, and the latter one represents literary texts with expressive features, which is rarely seen being tackled in crowdsourcing environments. My observations collected empirical data which examined the 'How' and 'Quality' aspects through real crowdsourcing translation practices.

8.1 Summary of results

In order to address the multi-dimensional perspectives of the research questions, a component of my project was to investigate the **profile, self-declared motivations, mechanisms for participation,** and **quality control methods** drawn from the questionnaire survey. **Chapter 4** detailed these findings.

In terms of the demographic profile of crowdsourcing translation participants,

my research identified that people with a higher education background and L2 competence are more likely to contribute. The crowd is composed of non-professionals (amateurs), amateurs who received translation-related training, and professional translators. Moreover, a positive trend was discovered between the experience of the respondents and the roles they played in the community: the higher the participants' L2 competence is, the more likely they are to be involved in influential roles, such as project manager (PM), reviser, and translator. Interestingly, the role of 'reviser' is more likely to be taken by professional translators than those whose current occupation is *student*.

8.1.1 Motivations

The first objective of my research was to investigate why the crowd volunteer to contribute to crowdsourcing translation which has a cost in terms of their own effort and time. After closely examining the current research in motivational studies, I chose the frameworks of Cognitive Evaluation Theory, Self Determination Theory, and General Interest Theory in order to analyse the motivating factors.

My assumption was that the motivations that drive people to contribute to *Yeeyan* largely depend on intrinsic factors given that most of the activities are carried out for free or rewarded with low payment (as in the *Gutenberg Project*). However, my investigation revealed that people are usually driven by multiple reasons both intrinsically and extrinsically. Intrinsic motivation implies that people perform an activity that can fulfil their psychological needs: the need for autonomy, the need for competence, and the need for social-relatedness. Extrinsic motivation implies that people perform an activity to receive external benefits, accumulation of personal skills, reputation or other forms of rewards.

- the **intrinsic motivations** include: (1) altruism, pursuit of (2) community identification, (3) enjoyment, and (4) self-achievement;

- the **extrinsic motivations** are rewards through (1) the enhancement of human capital, (2) self-marketing, and (3) peer recognition, as well as (4) personal needs.

According to the respondents' ranking of the motivating factors, my research identified that **extrinsic reasons** were perceived by the crowdsourcing translators' community I analysed to be **more important than the intrinsic reasons** (see the average rank of motivations in Table 9, section 4.2.3). This contradicted both my initial assumption and previous studies on this topic (Dolmaya, 2012; Olohan, 2012).

The participants in my study were more motivated by the chance to 'improve language abilities' and 'gain translation experience' to increase their *human capital*. This is a form of future rewards through the accumulation of personal skills, capabilities, and knowledge which can lead to better job opportunities and more fulfilling jobs. It fits the context of Chinese crowdsourcing translation initiatives: trainee translators and students in language-related majors are the main labour force as they use the crowdsourcing platform as a practice-ground to improve their skills and capabilities and to prepare themselves for the multifaceted challenges in professional Language Services Industry. 1/3 of the respondents in the questionnaire survey were students and selected the improvement of language and translation abilities as their primary reason for participation. The self-declared benefits which participants had gained through contributing to translation-related activities in *Yeeyan* proved the crowdsourcing platform and activity per se's potential to meet their expectations, which enabled the achievement of personal goals that motivated them.

In contrast to Sheldon and Elliot's (1998) and Stewart et al.'s, (2010) research in which intrinsically-motivated goals are associated with the most effortful behaviours, as compared to extrinsically-motivated goals, my study found through the analysis of motivations and the time one spends on contributing to translation-related activities in *Yeeyan* that people who are primarily motivated by extrinsic reasons are likely to invest more time and effort in crowdsourcing translation than those motivated by intrinsic reasons.

69% of the respondents to my survey perceived the increase of human capital as *the most important* motivation and spent an average of nearly 11 hours per week translating in the community, which is 2.5 times than those who see enjoyment as the *most important* motivation.

In addition to the aforementioned factors that positively affect the crowd's participation, my research also identified a number of **de-motivators** which the participants in my study expressed in the self-reflection section of my survey.

On the micro level, the settings and properties of the translation projects were mentioned, including:

- (1) the topic, length and degree of difficulty of the source text,
- (2) the project schedule and rules,
- (3) the competence of the team members.

These factors can negatively influence the crowd's decision to take part in the crowdsourcing translation projects or not. They can be correspondingly interpreted as:

- (1) limiting the perception of enjoyment and self-achievement,
- (2) hindering the experience of autonomous participation,
- (3) weakening the opportunity to meet the expectations set by the contributors, such as improvement of human capital and gains in reputation.

On the macro level, my research also indicated that the environment in which the translation activity takes place is highly valued by the respondents, who expressed concerns regarding:

- (1) the communication between members,
- (2) the feedback mechanism,
- (3) the evaluation metrics for the individual's contribution,
- (4) the user-friendliness of the crowdsourcing platform.

A negative performance of the first two factors will not meet the crowd's need for social relatedness and will also hamper the experience of reciprocity and

self-efficacy. Moreover, an unsatisfactory implementation of the third factor – the evaluation metric – will diminish the opportunity for self-marketing and pursuing peer recognition. Finally, the fourth factor directly affects the experience of executing translation-related tasks – i.e. the user-friendliness of the text editor and collaboration platform determines whether the participants can contribute comfortably and productively, which further influences the perception of enjoyment.

8.1.2 Participation

The second objective of my research was to explore how an individual's contribution is valued in the crowdsourced environment. The full-functioning of the *Yeeyan* community is supported by diverse forms of participation. The respondents to my survey stated the roles they had fulfilled in the community, which ranged from source text provider to PM, translator, reviser, mentor, reader, commentator, and promoter. My research discovered that, unlike more traditional, professional translation projects where roles are fixed and project participants do not usually fulfil more than one role, an individual taking part in a crowdsourcing project usually plays multiple roles during the content circulation and generation process.

The aggregation of contributions through these roles highlights a major difference between professional and crowdsourced projects, and indicates how the collective effort supports crowdsourcing translation on a general level.

The empirical data I collected through **observation** tracked the individual's performance in collaborative translation project, and demonstrated the **natural role-switching** in real practices. It also highlighted the self-organisation and autonomous features in *Yeeyan* as a community-based crowdsourcing initiative in which the high accessibility of peer contributions is essential (i.e. project members can view, comment and edit peers' work).

Workflow mapping in **Chapter 5** visualised the full extent of the participants' contribution to promoting the progress of collaborative translation projects. This complemented the results of the questionnaire survey which separately

highlighted **the sense of responsibility and ethics** felt among the *Yeeyan* members when involved in the translation and its related activities, particularly for those with strong extrinsic motivations. Respondents who value and abide by the ethics codes in professional conduct were primarily motivated by extrinsic reasons (i.e. the enhancement of human capital). Such a characteristic was empirically reflected in the real scenarios: when individuals volunteer to take part in translation projects, a majority of them spontaneously switch their role from translator to reviser (performing self-revision and peer-revision) according to the progress of the project even without an explicit schedule design and concrete revision instructions. Therefore, this research further challenged the criticism regarding the crowd being too difficult to control. Participants in my study demonstrated being able to self-organise and collaborate spontaneously, as well as perform translation-related activities to a high standard.

'Workflow and Productivity'

One noteworthy finding of my research is that revision takes much longer than translation in the crowdsourced environment, which adds a new perspective to the widely-claimed advantage of crowdsourcing translation to produce a *high* volume of translation with a *low* cost and at a *fast* speed. My research findings contradict in fact the 'speed' aspect: the workflow mapping and **productivity analysis** (section 5.2.1) demonstrated that the participants' 'translation' speed for a large volume of text is fast, but if quality is also taken into consideration, **the crowd needs a much longer time to perfect their translation** because the crowd's revision productivity is much lower than the expected professional industry standard which also applies to translation trainees.

Furthermore, because my quality assessment (discussed in full in **6.4**) of all the observed translation outputs proved that peer-revision is the most effective approach to improving quality, my research found that an individual's engagement in the revision process can be used as a metric to weigh one's overall project contribution.

Two kinds of workflows were identified in which the individuals' contribution is

weighted differently. In the four *General Collaborative Projects* I examined whose main aim was knowledge dissemination, I observed that a **user-driven flexible workflow** was adopted, which mainly depended on the participants' spontaneity which, in turn, guaranteed the participants' autonomy to a large extent. A **network organisational structure** was embedded in this workflow and, under such conditions, I observed that the primary participants' contribution was distributed **unevenly**, which was also mirrored in the effort that they invested in the revision activity. I was able to identify **lead participants** among those who are passionate about collaborating with others and who consistently put more effort in terms of *quantity* and revising peers' work in particular.

In the *Gutenberg Project* whose output was aimed for publication, a different, **waterfall workflow** was implemented. The progress of the project was strictly managed and monitored by the PM, who allowed participants less autonomy and implemented a much tighter **hierarchical organisational structure**. The two primary participants' contributions were **equally** distributed, i.e. they commenced a similar number of tasks. Another difference between the two types of workflows and organisational structures lies in the significance of the PM's contribution: my research found that the **PM's role** was more important in the waterfall workflow than in the user-driven flexible workflow. Both professional-like project management skills and good linguistic capacity were displayed by the PM in the *Gutenberg Project.*

'Dialogue Acts Analysis and Social Network Analysis'

Online peer-to-peer interaction is a prominent characteristic of collaborative translation projects in *Yeeyan*. The individuals' contributions were manifested in the conversational discourse throughout the projects (see **Chapter 6**), which I was able to annotate and evaluate using **Dialogue Act Analysis** as a framework. I classified instances of **turn-taking** present through discourse units such as questioning, clarifying, discussing, and accepting. In addition to translation-related roles, I also observed that participants in the crowdsourced environment contributed in new positions – e.g. as information provider, information requester, problem-solver, coordinator, and activator.

The ongoing structure of turn-taking sequences proved to be an indication of the degree to which information sharing and knowledge exchange influences the *progress* and *quality* of collaborative projects.

Studies in crowdsourcing have emphasised the importance of the leading user in promoting participation and quality control. Although participants in the crowdsourced environment, as described in a Participatory Culture, tend to self-organise the process, my initial hypothesis was that there should be a key user who actively performs various activities, coordinates the participants and promotes the development of the project. Using **Social Network Analysis**, I examined the information flows and the relationships between the participants connected by interactions in the observed projects. I was thus able to identify the **key actors** in the interaction networks by using a combination of the '*weighted out-degree centrality'* metric, specialised software, and visualisation techniques. I found that the key actors made extensive effort to stimulate the peers' engagement and help establish a cooperative pattern in which information flows smoothly. More importantly, their contributions as information providers and problem-solvers were significant in quality control mechanisms such as the **feedback** loop which was formed in the interaction network. The key actors took the initiative of providing 'revision comments', which embodied two kinds of peer feedback: (1) simple feedback, offering an evaluation of one's translation, informing whether the translation is good or problematic, and explaining what a good translation would be; and (2) elaborate feedback, explicitly pointing out where the error is and how it can be improved, as well as solving peers' translation-related problems through the dialogue acts of clarifying and discussing.

The identity of '**key actors**' in the social network analysis corresponds to that of the '**lead participants**' in the workflow mapping. This finding highlighted the importance of involving such contributors in crowdsourcing translation projects, because in addition to the initially-assigned task, they are **more willing to** engage in a wide range of additional activities, **investing more effort and time in linguistic tasks** (i.e. translation and revision) and **leading non-linguistic activity** (e.g. peer-to-peer interaction), which all drive the project to in a positive direction.

Compared with the professional Language Services Industry, an individual's contribution in a crowdsourcing project appeared to be valued much more diversely, which fits with the core principle of Participatory Culture which states that 'everyone's contribution matters'. It is the aggregation of fractions of individual effort in translation-related activities and peer-to-peer interaction that pushes collaborative projects to advance in the crowdsourced environment. However, I have found that it was the lead or key participants who put more effort into improving their peers' translations, and promoted the knowledge-sharing in peer-to-peer interactions. These participants are the indispensable contributors that lead to the success of the crowdsourcing translation projects, and they should be clearly identified and acknowledged, as well as safely trusted with more responsibilities and tasks.

In addition, I also conducted a comparative study between the crowdsourced environment and the University of Leeds Centre for Translation Studies *Students' Team Project* which mirrors multilingual localisation projects in the professional industry (section 6.3, **Chapter 6**). I found that the identity of key actors in the interaction networks in the two environments is different: in the *General Collaborative Projects* characterised by a network organisational structure, the primary participants were the translators, whereas in the *Gutenberg Project* and the *Students' Team Project* which featured a hierarchical organisational structure, the PMs' contribution was more significant than other participants' (or freelancers'). This finding highlighted the different power relationship between the project actors, and provided evidence on how the individuals' contribution was valued in these two different organisational structures.

8.1.3 Quality assessment and quality control

The third objective of my research was to investigate the quality of crowdsourcing translation. For this purpose, I conducted an **Error-based Quality Assessment** to examine three versions of the translation outputs in each crowdsourcing project: the first version of the translation, the subsequent self-revised version, and the final peer-revised version (see **6.4**). The error categories which I used in the annotation stage came from a

modified version of the *MeLLANGE* typology.

The empirical data I acquired and analysed indicated that participants in the crowdsourced environment are generally only capable of producing draft-level quality that needs vigorous revision, but eventually, final '**fit for purpose**' translations can be produced through collaborative workflows. My analysis also built on professional industry criteria in the Chinese market (as stated in *GB/T 19682-2005, GB/T19363.1-2008*). The 'fit for purpose' result was a consequence of the high accessibility of peer contributions in that particular environment. The statistical analysis I performed on the error distribution between the first draft and the final version showed that **more than 95%** of errors on average were corrected through the collective effort of the crowd in the observed projects.

Among the quality control methods used in crowdsourcing projects, the respondents to my questionnaire survey mentioned:

- the preselection of contributors,
- working with a shared glossary list,
- discussing frequently with team members,
- involving a subject-matter professional or PM to do the final review,
- performing self-revision,
- performing peer revision.

I used empirical data to examine the effectiveness of these self-declared quality control methods. First of all, the error annotation stage proved that having a shared glossary is useful for reducing inconsistencies and improving content accuracy. Also, my annotation of the communication data highlighted the importance of the specific dialogue act of reminding translators to add the specialised terms and names to the glossary list. Participants who performed the revision activity relied on this glossary list.

Peer-to-peer interaction helped eliminate **11%**, **12%**, **3%**, **8%**, and **4%** of the errors identified in the first draft (in *Health Project I*, *Health Project II*, *Business Project I*, *Business Project II*, and the *Gutenberg Project*, respectively). The rest of the errors were spotted and corrected through peer-

revision. Lexical issues were the most frequent topic in dialogue acts that directly related to the substance of the translation. Generally, the rate of improving translation quality through communication was relatively low, yet I observed that participants often turned to the chat board for solutions whenever they could not solve a translation problem on their own.

The preselection of contributors was a quality control method exclusively applied in the *Gutenberg Project* for quality reasons. My research showed that a sample test is effective for selecting contributors with higher translation competence, and is also particularly useful for filtering out those with lower target language capabilities (compared with the untested participants who signed up to the *General Collaborative Projects* freely, the preselected translators made significantly fewer errors in AW (awkward syntax) and NN (not idiomatic and non-native use of target language), an achievement of particular importance in a literary project that aims for publication). In the case of literary translation, accurately conveying the source content is not the only quality criterion, and the ability to render the translation naturally in accordance with the rules and conventions in the target language is highly valued, as well. My research therefore confirmed the validity of using a sample test to preselect contributors with better translation competences among the crowd, especially for literary texts that require enhanced skills in processing the target language.

Regarding the revision procedures, the data I gathered showed that in the crowdsourcing community I examined, **peer-revision** was the most efficient approach to improving translation quality. A T-Test revealed that there was no significant reduction of the error counts in nearly all error categories after self-revision, with the exception of term error (TE), where the p-value was less than 0.05. However, despite being able to correct one's own errors, a correlation analysis identified a positive relationship between the errors that one makes and errors that one corrects in peers' translations. The result indicated that **participants may not perform well in spotting and correcting their own mistakes in the self-revision stage, but they are capable of finding others' errors and making suitable revisions**, which explains why peer-revision was the most efficient quality control method.

My project showed that participants were particularly good at improving the accuracy of the translation by reducing *Content* errors at the lexical level (TE – an error rate of 26% and a correction rate of 97% on average), as well as *Language* errors covering non-idiomatic, non-native use of target language (NN – an error rate of 25% and a correction rate of 97% on average), and ambiguous translation (AM – an error rate of 8% and a correction rate of 94% on average).

Moreover, I found that the final revision conducted by the PM in the *Gutenberg Project* was essential for improving the *fluency* of the translation, as the PM's corrections were very comprehensive, covering all types of language errors, style errors (literal translation – LT), and hygiene errors (incorrect punctuation – PU and character spelling – CS). This further proved the importance of involving a third-party reviser – who could be an expert reviewer or a PM with good linguistic ability – in crowdsourcing translation projects with higher quality requirements.

An additional finding of the quality assessment I carried out in my project was the analysis of '**significant errors**' which indicated the weaknesses of the participants' translation ability. TE (26% on average) and NN (25% on average) were the top two error types made in the translation stage, and were also the ones being most significantly corrected in the revision stage. TE and NN belong to the two broad error categories: '*Content*' which caters for the accuracy of translation, and '*Language*' which deals with fluency of the translation. This finding is significant especially for projects in which the crowd is only recruited to perform the translation. Both participants and organisers of such projects need to be aware of these major error categories and plan modifications to the workflow to address them.

8.1.4 Text types that can be feasibly translated through crowdsourcing

The **fourth objective** of my research was to examine the **feasibility of processing different text types** in crowdsourcing translation projects. The quality assessment I performed confirmed that the translation of both

specialised and literary texts is feasible through crowdsourcing because the collaborative models can guarantee the translation quality.

One further notable finding concerned the differences in the significance of error types which pointed out the obstacles in translating these two text types in the crowdsourced environment. Terminology is the biggest challenge for specialised texts, while *Language* errors occurred more frequently in the translation of literary text as participants had difficulty in rendering the translation fluently and elegantly. However, these problems can be dealt with successfully through collective effort.

Moreover, I also identified the subcategories of the significant error types in the two text types (details see section 7.3.2) – for instance, within the group of terminology errors (TE), specialised terminology is more evident in specialised texts (38% in Business, 36% in Health and 1% in Literary), while collocation (32% in Literary, 16% in Health and 18% in Business) and contextual adaptation (31% in Literary, 22% in Health and 18% in Business) are more prominent in literary translations. In terms of *Language* errors (NN), missing function word, word order and redundant expression are more significant in the specialised texts. On the other hand, complex sentence structure which leads to misunderstandings and influences the fluency of reading is more frequent in literary translations. My results pointed out the specific aspects that need special attention in the future organisation of crowdsourcing projects with different text types and contributors selected according to differing criteria.

8.1.5 The learning perspective

The **final objective** of my project was to explore **the learning perspective** associated with participating in crowdsourcing translation initiatives. The respondents to my questionnaire survey and the participants in the *Gutenberg Project* acknowledged the acquisition of knowledge and skills through self-assessment, naming the improvement in language and translation ability as the most important benefit for them. Learning was achieved on three main levels:

- (1) the individual fulfilled multiple roles through the activities in the crowdsourcing community which diversified one's experience,
- (2) the knowledge exchange through peer-to-peer interaction,
- (3) elaborate feedback reflected in the revision edits.

The integrative contribution setting and the high accessibility of peer contributions embedded in collaborative workflows are the core components that support the learning process. Crowdsourcing translation provides an authentic industry-informed environment for participants to practise their desired skills, and the wide range of text types available in *Yeeyan* give participants the opportunity to experience various functioning roles, the environment to raise numerous questions, practise problem-solving through collaboration, as well as the reap the benefits of continuous individual feedback. Such advantages are not always easy to embed in formal classroom settings because of limited resources and tight schedules.

Furthermore, through social network analysis, I observed diverse interaction patterns (*one-to-many, many-to-one,* and *many-to-many*) in the crowdsourced environment which are beneficial for collaborative learning, as well. As opposed to the traditional formal training environment, in which hierarchical relationships dominated the content and direction of information exchange, the diverse interaction patterns in crowdsourcing projects allow the participants to decide what information they want to receive, to designate the person who provides feedback, and to debate controversial points. The exclusive *many-to-many* interaction pattern present in this environment maximised the coverage of knowledge exchange among participants. Moreover, the 'key' actors in the interaction networks resemble the role of instructor in the formal training environment – because they are actively involved in the translation project, they are more sensitive to the peers' needs, and can find more effective ways of giving explicit feedback to designated participants without de-motivating them.

Consequently, my research suggested that crowdsourcing translation environments can also be used very effectively as an alternative to traditional learning settings.

8.2 Suggestions

Based on the empirically-proven findings of the study, I propose the following suggestions that will help promote more effective organisation of crowdsourcing translation projects and result in better crowdsourcing outputs.

- **Motivation mechanisms**

Yeeyan is a community-based crowdsourcing initiative that relies on the members' shared interests and preferences. My investigation has highlighted how motivations and other factors influence participation of crowdsourcing translation, and therefore I recommend that crowdsourcing initiatives like *Yeeyan* take into account several factors that are likely to affect the crowd's motivations. It is important for the crowdsourcer to understand why people volunteer to contribute so that customised motivation mechanisms can be established. For instance, an *open collaborative model* implemented in the observed *General Collaborative Projects* would appeal much more to contributors who need social-relatedness, autonomy, as well as the enrichment of human capital and reputation.

Another possible option is to offer *flexible approaches* for the crowd to choose the desired projects and tasks, which will cater to individual preferences. For example, 88% of the respondents to my survey mentioned that they choose whether to participate in a crowdsourcing translation project depending on the topic of the source text, as well as the roles available to them. Therefore, categorising the source text in accordance to its topic instead of presenting everything scattered in the repository can attract a group of contributors with a shared interest, which may reinforce their collective power and perception of enjoyment. Moreover, it will also increase the chance of involving *lead participants* and *key actors* who are found to improve crowdsourcing translation from both quantitative and qualitative perspectives.

- **A thorough user profile system**

My study identified a network organisational structure in the crowdsourced environment in which no formal employment relationship is agreed between the crowdsourcer and the labour force, making the crowd highly mobile. Consequently, I suggest that crowdsourcers build a complete user profile system which can help maintain and consolidate the user community.

To be more specific, keeping track of user behaviour aggregates metadata (e.g. which topic the person is interested in, how much time one invests in translation-related activities in the community, and the roles one has fulfilled in the community), which enables the crowdsourcer to be sensitive and adaptive to user needs and preferences. Moreover, a continuous evaluation of the participants' performance will help devise realistic and engaging tasks for the crowd. This will be particularly useful for outsourcing portals that offer crowdsourcing translation services, as well as additional community-based crowdsourcing initiatives. Automatic notifications can be enabled to push tasks to the existing users that meet criteria of interest and competency through the computing of the metadata. Alternatively, the user profile can be used as reference for the manual preselection of participants when specific skills or attributes are desired, and it may also provide insights into establishing up-to-date motivation mechanisms.

- **Built-in interactive functionalities**

One obvious reason for having an interactive environment on the crowdsourcing platform or within the crowdsourcing community is to strengthen the relationships among the crowd. Social interaction is beneficial in mobilising the participants' engagement in crowdsourcing translation activities, and it also provides a space for contributors to gain feedback from both their peers and their audience. The interaction typology in my research highlighted the complex nature of the crowd's interactions by using thematically-grouped dialogue acts.

Moreover, the questionnaire survey in my study has showed that 'communication' is an important feature anticipated by the crowd for the

purpose of collective problem-solving, which helps improve translation quality. Although crowdsourcing participants tend to overestimate the effectiveness of collaboration for self-revision purposes, my empirical data has proved that peer-to-peer interaction is not only useful in problem-solving but also in coordinating and boosting individuals' contributions to the project. The benefit of having a built-in interactive functionality is particularly highlighted in the organisation of crowdsourcing translation projects. In cases like the *General Collaborative Projects* in *Yeeyan*, an interactive space extends the degree of self-organisation: contributors can use the interactive function to perform activities that project managers would usually coordinate. Such interaction helps advance projects in a positive direction through dialogue acts such as 'commission', 'revision request', 'motivation' and 'status check'.

In my research, I integrated an interaction typology to measure the benefits of collaborative learning and project management. *MNH-TT* is the only collaborative platform to date that embedded thematically-grouped dialogue acts to facilitate translation training (Babych et al., 2012). Building on *MNH-TT*, I suggest encouraging the participants to select each time the appropriate act type to classify the messages they send. When the project manager's involvement is limited, retrieving the back-end data on the frequency and categories of dialogue acts on a regular basis can provide a reliable measure regarding the progress of the projects and the degree of the individuals' contributions (more details in **Chapter 6**). For instance, if the data signals a large number of discussions regarding term issues ('term request' and the subsequent 'comment') and few signs of agreement ('term request' and the subsequent 'acknowledgement'), the crowdsourcer may need to consider inviting a terminologist to the project. This approach can be used as a convenient method to manage and monitor crowdsourcing translation projects, which enables the crowdsourcer to detect potential risks in advance and to implement flexible solutions.

Another advantage is the possibility to use the interactive data as a metric to assess the individuals' contribution during the project. Keeping track of the kinds of interactions can inform the design of reward mechanisms for

crowdsourcing initiatives, such as an upgrading system. As applied in my research, both *dialogue act analysis* and *social network analysis* can be used to identify the key contributors based on the participants' performance in group interaction. For instance, those who put great effort in mentoring, offering suggestions for problem-solving, or providing feedback can be acknowledged as key actors or high-level participants. Combined with their user profile data, more responsibility may be entrusted to these contributors, i.e. assigning the significant roles of project manager, reviser, or even assign them paid work.

- **Emphasising the 'revision' stage in a TEP-like workflow**

The workflow mapping, dialogue act analysis, and translation quality assessment in my study all emphasised **the significance of the 'revision' stage** in a crowdsourcing translation project. To begin with, because the crowd's productivity regarding revision is much lower than the expected standard in the professional industry, and also because the individuals' participation in revision is less structured than in translation, it takes much longer to have all translation outputs revised than merely performing the translation task. For crowdsourcing initiatives using TEP-like workflows, I suggest the crowdsourcer or project manager to reconsider the design of the schedule, allowing more time for revision and encouraging the participants to centralise the revision efforts so that the schedule is not delayed. For instance, the *Gutenberg Project* in this study exceeded the initial deadline by a whole six months because of the repeated delays in the revision stage. More importantly, the crowdsourcer should be aware that a TEP-like workflow may be challenging for translation projects whose priority is 'speed', but it will work effectively for projects which emphasise 'quality'.

Regarding the two types of workflows identified in this study (the *user-driven flexible workflow* and the *waterfall workflow*), they all led to successful projects in the crowdsourced environment. However, a certain degree of uncertainty is embedded in the former approach: on the one hand, it is completely influenced by the network organisational structure in which the relationship between the crowdsourcer and the labour force is not secured;

on the other hand, not all individuals are willing to invest time and effort in revising others' work. In the observed projects, the lead participants were the ones who contributed the most during the revision stage. At the same time, they were also those who coordinated and encouraged other individuals' engagement with the revision stage in the interaction network. It is therefore absolutely necessary to involve enthusiastic and good performers in the revision and management of such crowdsourcing translation projects. My research has provided guidance on how good contributors can be identified among the crowd.

Furthermore, this study proved that 'more involvement' is better than 'less involvement' in the sense that the crowd is better at spotting and correcting others' mistakes than correcting its own; the crowd also showed a low likelihood of introducing new mistakes. Consequently, I suggest emphasising the revision procedure in the crowdsourced environment by increasing the accessibility of peer contributions for quality assurance purposes. A motivating mechanism that inspires the crowd to critically comment on and revise others' work is beneficial to crowdsourcing translation projects for 'bigger tasks' (e.g. large source texts).

On the other hand, my research suggested that the 'waterfall workflow' is suitable for crowdsourcing initiatives in which a formal relationship between the crowdsourcer and the labour force can be secured – e.g. a success-based monetary reward. This is possible when the crowd is not concerned with their 'autonomy' during participation, which happens if crowdsourcing translation projects are set in a *closed* model like the *Gutenberg Project* in this study, involving high-level participants (e.g. a professional reviewer or project manager with high linguistic ability) whose presence is essential to ensure the crowdsourcing quality, especially when quality is the main project priority.

- **Structured feedback mechanism**

To improve the crowd's translation competence as a whole, I propose to implement a common, consistent error-based feedback mechanism in crowdsourcing translation projects that supports the accessibility of peer

corrections. This can be used in various scenarios, including a TEP-like approach (e.g. *Yeeyan*) or a participatory mechanism in which audiences can edit or report translation issues (e.g. *Wikipedia* and *FOSS*). By asking the contributors or the audiences to select an error type each time they edit, revise or critically comment on crowdsourcing translation outputs, quality loops can be thus formed for the benefit of both the crowdsourcer and the participants.

In terms of the contributors, given that the primary motivation for participating in crowdsourcing translation in *Yeeyan* is to improve translation and language ability, an error-based feedback mechanism can be a powerful motivator. Compared with the feedback one receives through the interactive functionality which may be informal, contingent, and subjective, my proposed feedback mechanism with a robust error typology would offer structured, stable and objective feedback as long as revision edits were applied. It will help contributors better understand why the translation needs revision and to offer elaborate feedback regarding the skills which each individual needs to improve. The aggregation of long-term error-based feedback may have a similar influence as an instructor-led training environment. If the crowdsourcing initiative is associated with formal translation training, a systematic error-based feedback overview will provide longitudinal evidence that illustrates the contributors' development after a series of participations (to be investigated in future studies).

As for the crowdsourcer, a regular review of the significant error types that the participants make will highlight any collective weaknesses, and will inform further customised training and supporting resources to reduce the observed error types, such as guidebooks on the use of punctuation, or specifications on the standard units of measure and foreign names.

In my study, specialised terminology (in the *General Collaborative Projects* with specialised texts) and target-language collocations (the *Gutenberg Project* with a literary text) were the common mistakes observed among the participants. In such circumstances the crowdsourcer may consider providing training on using language corpora and similar online resources to solve

lexical issues. It is a straightforward approach to train the crowd for a better outcome when taking a long-term strategic approach. Moreover, the feedback mechanism is a cost-effective and time-saving way to assess individuals' contributions, as well. In short, a structured feedback mechanism will enhance the recruitment and retention of the labour force for crowdsourcing initiatives, as well as the resulting quality outputs.

- **Set a 'difficulty score' of the source text to fit the individuals' competence**

When applied to a crowdsourced environment, the TEP-like workflow controls translation quality by reallocating the responsibility for the successful completion of various stages in the translation process. It empowers the crowd for the execution of translation-related activities, and in some cases, invites gatekeepers (high-level participants) to bear most of the responsibility for ensuring the quality of crowdsourcing outputs. My new suggestion is a *universal* approach even if a TEP-like workflow or gatekeeper is absent from the projects.

For large source texts, I propose to set a 'difficulty score' to fit the diversity of the crowd's translation competence. Adding the difficulty score to the project description will both manage the participants' expectations and make the crowdsourcing process more efficient, as I have found that 30% of the respondents to my project questionnaire consider 'the source text is too difficult' as a de-motivating factor. The reason for defining a difficulty score is to give additional relevant information to potential participants before joining a crowdsourcing translation project. In addition to using 'the topic of the source text' as a measure to decide whether to contribute, the difficulty score will also help participants select a project and sub-task that matches their ability.

Having a difficulty score is likely to reduce the withdrawal rate because participants would be better prepared psychologically for the project. Moreover, by adopting an 'open' entrance approach (rather than a preselection test), the participants still have the autonomy of deciding which project to join. It would be an effective approach to align contributors with translation projects that match their ability, which will subsequently help

prevent crowdsourcing output of extremely low quality. More importantly, this may also reduce the time, effort and cost of the revision stage.

The crowdsourcer or project manager can use term extraction tools at the beginning of the project to compile an initial glossary; the density of specialised terms will represent a metric to define the ‘difficulty score’ of the project: the higher the number of specialised terms, the more difficult the translation project is expected to be. If the crowdsourcer prepares a glossary list based on the term extraction result, it also provides *anticipatory feedback* to the crowd this way.

In my research, I used *Sketch Engine* (a corpus linguistics tool) to pre-analyse the source text (performing a *keyword* extraction) and to define the difficulty level in combination with computing a *reading score*. Commercial computer-assisted translation tools, such as *SDL MultiTerm* and *MemoQ*, as well as freeware, such as *AntConc*, *TaaS*, and *Lexterm* support term extraction, too. For a detailed discussion on how term extraction works, see Xu and Sharoff (2014), which suggests how to integrate term extraction into the preparation workflow.

Moreover, the difficulty score would also function as a *gamification* motivating mechanism among the crowd, which is found to positively influence in the participation of crowdsourcing activities (Morschheuser et al., 2016). The difficulty score will allow the participants to track the tangible progress of their translation ability in time. High-level participants and entry-level participants can also be naturally integrated through the mechanism, as well.

In summary, the proposed suggestions cover the retention and enhancement of the participants and their skills, the crowd motivating mechanisms, technological support, workflow design, and quality control methods which the findings of my research suggest will facilitate the efficient organisation of crowdsourcing translation projects and foster better crowdsourcing translation quality.

8.3 Contributions

This study provides a full-scale examination of crowdsourcing translation from **multiple dimensions**. It is one of the very few studies that investigates not only the user perspective (motivations and participation), but also the implementation and governance issues from the organisational perspective.

To begin with, the combination of the user perspective and the organisational perspective in my study **bridged a notable gap in the translation process and translation quality:** how translation evolves though aggregated contributions. My research analysed empirical data to highlight that the final quality in collaborative translation projects is influenced by multiple factors: the participants' translation competence is not the only variable, because quality is also influenced by the workflow, the degree of collaboration, and occasionally the influence of the project management approach, too. I also proposed **suggestions** for enhancing the collaboration between participants and implementing quality control methods for better crowdsourcing outputs.

Furthermore, my study highlighted **the influence of 'revision' procedures on the quality of the translation**. According to the quality standards in the translation service industry (i.e. ISO 17100:2015), revision (bilingual editing) should be included as a minimum in the translation production process. It is clear that 'revision' is highly valued in the professional industry, whereas the importance of 'revision' is often underestimated in translation training and underrepresented in curriculum design. My data suggested that participants in the crowdsourced environment perform better when revising others' translations compared to when revising their own work. However, in the professional industry, most of the translation work depends on self-revision for quality assurance, which is clearly problematic. The question is how to prepare trainee translators to perform effective revision. Using empirical data, this study presented the rationale together with several suggestions for involving 'revision' as a formal aspect of translation training. For instance, I identified the significant error types that should be addressed in the training environment. I also proposed to implement an error typology in the revision stage to provide structured and elaborate feedback to enhance learning, and

to generate longitudinal evidence to demonstrate the competence acquisition process. The error typology in my study is ready to use in any translation training environment and does not apply only to European languages. It can also be used in future studies to evaluate translation quality alongside individuals' performance.

My research also included a new aspect in Translation Studies: the influence of '**communication**' in collaborative translation projects. Scholars have demonstrated that crowdsourcing communities can be good sources for feedback to improve translation competence (Jiménez-Crespo, 2017; Orrego-Carmona, 2014). However, few studies had investigated the actual feedback process. I used empirical interaction data to map the giving and receiving of feedback through *Dialogue Act Analysis*, which proved the feasibility of using crowdsourcing for translation skills acquisition.

More importantly, I conducted a comparative study focusing on peer-to-peer interactions between the crowdsourced and industry-like environments (*Students' Team Project*) which revealed the interaction patterns in the two environments, and the extent of the individuals' contributions to the success of a translation project. These findings offered insights on how to leverage the collective intelligence and how to devise realistic and effective tasks in the two environments. My analysis of the detailed project logs, as well as the social networks occurring in the *Students' Team Project* highlighted the importance of different actors throughout a large translation/localisation project. It also acknowledged the significant effort and time put in by project managers, which is often overlooked in the Language Services Industry and translation training programmes. My results suggest that the average time a PM puts into a project can be up to **three times** more than the freelancer, depending on the size of the project and general competence of all actors involved. For communication only, the PMs' effort in the project I observed was more than **tenfold** that of each freelancer's. Therefore, this research empirically proves the **importance of integrating project management skills in translation training curricula**.

Moreover, my study also contributed to the introduction of **novel research**

methods in Translation Studies, namely *workflow mapping*, *dialogue act analysis* and *social network analysis*. In addition to the traditional textual analysis, the new methods bring more angles (i.e. an individual's role and contribution, the design of the workflow, the extent of project management activities and peer-to-peer interactions) to the examination of translation production, which can point out the key factors that contribute to the success of a project, and can therefore inform translation training, as well.

To be more specific, for educational programmes that adopt project-based learning, visualisations generated using the aforementioned methods enable students to understand better the workflow in large translation/localisation projects, specifically concerning how an individual's performance influences the project, as well as how project management works. Most importantly, these methods help spot the weaknesses of a project and offer insights into how to improve it. The advantages of using these research methods can also be reflected in a company environment, in which one wishes to test the effectiveness of the project design and the productivity of employees. The two interaction typologies (one for the crowdsourcing translation project and the other for the *Students' Team Project*) developed in this study are ready to use in crowdsourcing initiatives and collaborative learning platforms that wish to capture the interaction between the participants and the knowledge hidden in dialogue exchanges.

My study also has **practical contributions**. Starting from the user perspective, previous studies on the motivations for participating in crowdsourcing translation activities only focused on the pro-motivations. My study included the de-motivating factors in order to offer more thorough guidance in recruiting participants and design a more effective crowdsourcing environment.

In addition, I examined the validity of using collaborative projects in the crowdsourced environment and showed that the degree of collaboration can influence the quality of the translated output. My research shed light on how collaboration can be enhanced to ensure the success of a crowdsourcing translation project, e.g. by using built-in interactive functionalities, involving

key actors, enabling the accessibility of peer contributions, and implementing motivating mechanisms. I also made data-informed recommendations on how to assign responsibility to the crowd: for example, key actors should be identified through *workflow mapping* and *social network analysis* and should be recognised as high-level contributors who are capable of playing the significant roles of reviser and PM.

My study also verified the feasibility of translating specialised texts and literary texts through crowdsourcing, thus enhancing the application of crowdsourcing translation to a wider range of content, given that literary texts are rarely seen in the crowdsourced environment. With the high quality of the final output in mind, I further identified the significant error categories occurring when translating the two text types, a finding which is insightful and useful for implementing suitable quality control methods (as stated in section 8.2) and improving translation training programmes, as well.

Last but not least, through the comprehensive evaluation of the *Yeeyan* community, which is the earliest and largest crowdsourcing translation initiative in China, my study enriched the collection of translation practices in Translation Studies. Much of the research to date on crowdsourcing translation focused on its applications in the West – for instance, *Facebook*, *Wikipedia*, *TED*, the *Rosetta Foundation*, and *Amazon Mechanical Turk*. Crowdsourcing translation has been thriving for more than a decade in China, yet little attention has been paid to these specific practices in this participatory culture. Given that China is one of the largest markets for content consumption, it has a large, constantly-growing demand for translation services, and that the majority of the labour force in crowdsourcing initiatives are students in translation and language-related majors, another contribution of my study was the raised focus on crowdsourcing initiatives in the Chinese context with its cultural and social specificities. This diversified the translation focus in academia and helped promote the framework of Translation Studies towards a more general coverage.

8.4 Research limitations

Despite the encouraging outcomes, I am aware of the limitations of this study. Firstly, due to the inevitable constrains on the schedule and availability of resources, the scope of the PhD study is limited: the experiment only managed to investigate five crowdsourcing translation projects together with their contributors as research subjects. The findings are therefore not sufficient for making sweeping statements about crowdsourcing translation in general, and especially about crowdsourcing in a different, less co-operative culture. Moreover, the study looked only at successful projects; further work is needed to investigate failed crowdsourcing projects. Replication studies with more research subjects and the investigation of failed projects are needed to further confirm the patterns of participation, workflow, and translation quality.

In terms of the learning perspective of participating in crowdsourcing translation projects, the findings are mainly based on self-assessment data and the researcher's analysis of the conversational discourse and peer feedback manifested in revision edits. For unavoidable reasons, the current research lacks a longitudinal study that tracks the growth of the participants' translation ability. Moreover, an additional quantitative approach will be needed to measure the outcome of learning in the crowdsourced environment.

8.5 Directions for future work

In addition to the remedies for the current research limitations, several topics for future work have been identified. First, my proposed suggestions need to be tested in relation to the desired outcome, as well as in terms of their acceptability by the crowd.

The current research results apply only to collaborative translation projects in the crowdsourced environment. I plan to integrate a crowdsourcing initiative with formal translation training since the result would allow participants to experience multiple roles as demanded in the professional industry and to

translate (or revise) a diverse range of text types through serious ‘deliberate practice’. Crowdsourced environments like *Yeeyan* are suitable for students who are constantly striving to acquire and practise professional-level skills and productivity; unfortunately, the limitations imposed by contact time and class size often mean that university courses are not able to offer such practice. Students enrolled in Master programmes in Translation Studies will be preferred in future studies as they are available for longitudinal studies which will allow the researcher to quantify the outcome of learning and to test the effectiveness of the proposed suggestions.

The collaborative learning process is certainly worth investigating further in conjunction with other relevant theoretical frameworks, as well. For example, techniques borrowed from Cognitive Translation Studies such as eye-tracking and key-logging can be used to investigate the comprehension and production process, as well as to capture the knowledge acquisition process.

Finally, I plan to apply corpus studies methods to larger crowdsourcing translation outputs – particularly ones with structured error annotations – in order to investigate how different error types impact translation quality. These annotated error types can be used as language modelling features for machine learning (Yuan et al., 2016). Such a research strand will also be useful in the development of automatic error classification and estimation of translation quality.

All in all, a lot of work has gone into this project and I have thoroughly enjoyed the task of combining novel research methods, challenging traditional assumptions regarding crowdsourcing translation, and investigating in detail both the people and the processes involved in this field. I am excited by my findings, but even more by being part of a growing community of translation studies scholars passionate about investigating the many facets of crowdsourcing translations. The journey remains long and challenging, but its purpose is practical and beneficial for all readers of content and all those thirsty for knowledge, expanding their horizons, and embracing other points of view and ways of life. I look forward to crossing paths with you, dear reader, along the way, to learning about your own

findings and to hearing if and how my own work has helped you on your journey.

Bibliography

Aalst, W., Hee, K., 2004. Workflow Management: Models, Methods, and Systems. MIT Press.

Aikawa, T., Yamamoto, K., Isahara, H., 2012. The Impact of Crowdsourcing Post-editing with the Collaborative Translation Framework. In: *Proceedings of JapTAL 2012, Advances in Natural Language Processing*, Springer, Berlin, pp. 1–10.

Aitamurto, T., 2016. Crowdsourcing as a Knowledge-Search Method in Digital Journalism. *Digital Journalism*. Volume 4, Issue 2, pp. 280–297.

Al-Qinai, J., 2000. Translation Quality Assessment: Strategies, Parametres and Procedures. *Meta*. Volume 45, Issue. 3, September 2000, pp. 497-519.

Ambati, V., Vogel, S., Carbonell, J., 2010. Active Learning and Crowd-Sourcing for Machine Translation. In: *Proceedings of the Seventh International Conference on Language Resources and Evaluation*. pp. 2169–2174.

American Society for Testing and Materials, 2014. Standard Guide for Quality Assurance in Translation: ASTM F2575 – 14. [Online]. [Accessed on 26/11/17]. Available from https://www.astm.org/Standards/F2575.htm.

Anastasiou, D., Schäler, R., 2010. Translating Vital Information: Localisation, Internationalisation, and Globalisation. *Syn-Thèses Journal*. Issue 3, pp. 11–25.

Andrade, H., Valtcheva, A., 2009. Promoting Learning and Achievement through Self-Assessment. *Theory into Practice*. Volume 48, Issue 1, pp. 12–19.

Babych, B., Hartley, A., Kageura, K., Thomas, M., Utiyama, M., 2012. MNH-TT: A Collaborative Platform for Translator Training. In: *Translating and the Computer*, No. 34, 29-30 November 2012, London.

Baer, N., Nagle, L., 2011. Scaling and Strengthening a Closed Crowd: Lessons from Kiva's Evolving Model. Presented at *TAUS Roundtable on Collaborative Translation*, 10 October 2011, Santa Clara.

Becker, G., 1962. Investment in Human Capital: A Theoretical Analysis.

Journal of Political Economy. Volume 70, Issue 5, pp 9–49.

Bey, Y., Boitet, C., Kageura, K., 2006. The TRANSBey Prototype: An Online Collaborative Wikibased CAT Environment for Volunteer Translators. In: *Proceedings of the 3rd International Workshop on Language Resources for Translation Work, Research & Training*. May 2006. pp. 49–54.

Bogucki, Ł., 2009. Amateur Subtitling on the Internet. In: Cintas, J. and Anderman, G., eds., *Audiovisual Translation*. Palgrave Macmillan, London, pp. 49–57.

Borst, I., 2010. Understanding Crowdsourcing: Effects of Motivation and Rewards on Participation and Performance in Voluntary Online Activities. *ERIM Ph.D. Series Research in Management*. Erasmus Research Institute of Management.

Brabham, D., 2008a. Moving the Crowd at iStockphoto: The Composition of the Crowd and Motivations for participation in a crowdsourcing application. *First Monday.* Volume 13, Issue 6. [Online]. [Accessed on 02/11/17]. Available from http://firstmonday.org/article/view/2159/1969.

Brabham, D., 2008b. Crowdsourcing as a Model for Problem Solving: An Introduction and Cases. *Convergence.* Volume 14, Issue 1, pp. 75–90.

Brabham, D., 2012. Motivations for Participation in a Crowdsourcing Application to Improve Public Engagement in Transit Planning. *Journal of Applied Communication Research*. Volume 40, Issue 3, pp.307–328.

Brabham, D., 2013. Crowdsourcing. MIT Press, Cambridge, Massachusetts; London, England.

Brunette, L., 2000. Towards a Terminology for Translation Quality Assessment. *The Translator*. Volume 6, Issue 2, pp. 169-182.

Bruns, A., 2012. How long is a Tweet? Mapping dynamic conversation networks on Twitter using Gawk and Gephi. *Information, Communication & Society*. Volume 15, Issue 9, pp. 1323–1351.

Callison-Burch, C., 2009. Fast, Cheap, and Creative: Evaluating Translation Quality Using Amazon's Mechanical Turk. In: *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*.

Volume 1, pp. 286–295.

Cámara, L., 2015. Motivation to Collaboration in TED Open Translation Project. *International Journal of Web-Based Communication*. Volume 11, Issue 2, pp. 210–229.

Cao, Y., 2015. Crowdsourcing Translation in Contemporary China: An Inspiring Perspective of Translation in the Web 2.0 Age. *Meta*. Volume 60, Issue 2, pp. 316.

Chesbrough, H., 2003. The Era of Open Innovation. *MIT Sloan Management Review*. [Online]. [Accessed on 06/11/17]. Available from http://sloanreview.mit.edu/article/the-era-of-open-innovation/.

Chesterman, A., 2014. Translation Studies Forum: Universalism in Translation Studies. *Translation Studies*. Volume 7, Issue 1, pp. 82–90.

China National Committee for BTI Education, 2017. Announcement – List of National Universities offering Master and Bachelor's Degrees (全国翻译硕士及翻译本科办校名录). [Online]. [Accessed on 01/02/18]. Available from http://cnbti.gdufs.edu.cn/info/1006/1519.htm.

Clark, H., Schaefer, E., 1989. Contributing to Discourse. *Cognitive Science*. Volume 13, Issue 2, pp. 259–294.

Cnaan, R., Goldberg-Glen, R., 1991. Measuring Motivation to Volunteer in Human Services. *The Journal of Applied Behavioural Science*. Volume 27, Issue 3, pp. 269–284.

Cronin, M., 2010. The Translation Crowd. *Revista Tradumàtica: Tecnologies de la traducció*. No.8, pp. 1–7.

Dalecki, L., Lasorsa, D.L., Lewis, S.C. 2009. The News Readability Problem. *Journalism Practice*. Volume 3, Issue 1, pp.1-12.

Deci, E., 1972. Intrinsic Motivation, Extrinsic Reinforcement, and Inequity. *Journal of Personality and Social Psychology*. Vol 22(1), pp. 113-120.

Deci, E., Ryan, R., 2000. The 'What' and 'Why' of Goal Pursuits: Human Needs and the Self-determination of Behaviour. *Psychological Inquiry*. Volume 11, pp. 227–268.

Deci, E., Ryan, R., 2002. Handbook of Self-determination Research. University of Rochester Press. New York.

De Groot, A.M.B., 1997. The Cognitive Study of Translation and Interpretation: Three Approaches. In: Danks J.H. et al. eds., C*ognitive Processes in Translation and Interpretation*. Thousand Oaks CA: Sage Publications. pp. 25-56.

Delwiche, A., Henderson, J., 2013. Introduction: What is Participatory Culture?, In: Delwiche, A. and Henderson, J., eds., *The Participatory Cultures Handbook*. Routledge. New York & London. pp. 3-9.

DePalma, D., Kelly, N., 2008. Translation of, by, and for the People. *Common Sense Advisory.* [Online]. [Accessed on 06/11/17]. Available from http://www.commonsenseadvisory.com/AbstractView.aspx?ArticleID=889.

DePalma, D.A., Kelly, N., 2011. Project Management for Crowdsourced Translation: How User-translated Content Projects Work in Real Life. In: Dunne, K. and Dunne, E., eds., *Translation and Localisation Project Management: The Art of the Possible*. John Benjamins Publishing Company. Amsterdam. pp 379-408.

Deriemaeker, J., 2014. The Power of the Crowd: Assessing Crowd Translation Quality of Tourist Literature (MA Thesis). Universiteit Ghet. Belgium.

Désilets, A., 2007. Translation Wikified: How will Massive Online Collaboration Impact the World of Translation? In: *Proceedings of Translating and the Computer* (29). 29-30 November 2007. London.

Désilets, A., van der Meer, R., 2011. Co-creating a Repository of Best-practices for Collaborative Translation. *Linguistica Antverpiensia*. No. 10. pp. 27-45.

Desjardins, R., 2017. Online Social Media (OSM) and Translation. In: Desjardin, R., eds., *Translation and Social Media: in Theory, in Training and in Professional Practice*. Palgrave Macmillan. UK, pp. 13–33.

Dijck, J. van, 2009. Users like you? Theorising agency in user-generated content. *Media, Culture, & Society*. Volume 31, Issue 1, pp. 41–58.

Doan, A., Ramakrishnan, R., Halevy, A.Y., 2011. Crowdsourcing Systems on

the World-Wide Web. *Communications of the ACM*. Volume 54, Issue 4, pp. 86–96.

Dolmaya, J.M., 2012. Analysing the Crowdsourcing Model and Its Impact on Public Perceptions of Translation. *The Translator.* Volume 18, Issue 2, pp. 167–191.

Dolmaya, J.M., 2011. The Ethics of Crowdsourcing. *Linguistica Antverpiensia*. *New Series – Themes Translation Studies.* No. 10. pp 97-110.

Dombek, M., 2014. A Study into the Motivations of Internet Users Contributing to Translation Crowdsourcing: The Case of Polish Facebook User-translators. (Doctoral Thesis). Dublin City University. School of Applied Language and Intercultural Studies.

Domingo, D., Quandt, T., Heinonen, A., Paulussen, S., Singer, J.B., Vujnovic, M., 2008. Participatory Journalism Practices in the Media and Beyond. *Journalism Practice*. Volume 2, Issue 3, pp. 326–342.

Drugan, J., 2013. Quality in Professional Translation: Assessment and Improvement. Bloomsbury Academic.

Dunne, K.J., 2011. Managing the fourth dimension: Time and schedule in translation and localization projects. In: Dunne, K.J. and Dunne, E.S., eds., *Translation and Localisation Project Management: The Art of the Possible*. John Benjamins Publishing Company. Amsterdam

Dunne, K.J., Dunne, E.S. (Eds.), 2011. Translation and Localisation Project Management: The Art of the Possible. John Benjamins Publishing Company. Amsterdam.

Eisenberger, R., Pierce, W.D., Cameron, J., 1999. Effects of Reward on Intrinsic Motivation – Negative, Neutral and Positive: Comment on Deci, Koestner, and Ryan (1999). *Psychological Bulletin.* Volume 125, Issue 6. pp. 677-700.

ELIA (European Language Industry Association), EMT (European Masters in Translation network), EUATC (European Union of Associations of Translation Companies), GALA (Globalisation and Localisation Association), 2017. 2017 Language Industry Survey – Expectations and Concerns of the European Language Industry.

Ellis, D., 2009. A Case study in Community-Driven Translation of a Fast-Changing Website. In: Aykin, N., eds., *International Conference on Internationalisation, Design and Global Development*. Springer. Berlin, Heidelberg. pp. 236–244.

EMT Expert Groups, 2009. Competences for Professional translators, Experts in Multilingual and Multimedia Communication. [Online]. [Accessed on 26/11/17]. Available from https://ec.europa.eu/info/sites/info/files/emt_competences_translators_en.pdf.

Ericsson, K. A., 2000. Expertise in Interpreting: An Expert-Performance Perspective. *Interpreting*. Volume 5, Issue 2, pp.187-220.

Estellés-Arolas, E., González-Ladrón-de-Guevara, F., 2012. Towards an Integrated Crowdsourcing Definition. *Journal of Information Science*. Volume 38, Issue 2, pp. 189–200.

European Committee for Standards, 2006. BS EN-15038 European Quality Standard for Translation Service Providers. [Online]. [Accessed on 26/11/17]. Available from http://qualitystandard.bs.en-15038.com/.

Filip, D., 2012. Localisation for the Long Tail: Part 1. *MultiLingual*. Volume 23, Issue 7, pp. 51–55.

Fox, B., 2009. The quest for quality: perceptions and realities. Presented at *the CIUTI-Forum 2008: Enhancing Translation Quality: Ways, Means, Methods*, Peter Lang Bern, pp. 23–28.

Frey, B., Stutzer, A., 2002. Happiness and Economics: How the Economy and Institutions Affect Human Well-Being. Princeton University Press, Princeton and Oxford.

Fujita, A., Tanabe, K., Toyoshima, C., Yamamoto, M., Kageura, K., Hartely, A., 2017. Consistent Classification of Translation Revisions: A Case Study of English Japanese Student Translations. In: *Proceedings of the 11th Linguistic Annotation Workshop*. Association for Computational Linguistics. Valencia, Spain. April 3, 2017. pp. 57-66.

García, I., 2014. Training Quality Evaluators. *Revista Tradumàtica: tecnologies de la traducció*. No. 12, pp. 430–436.

Geiger, D., Seedorf, S., Schulze, T., Nickerson, R.C., Schader, M., 2011.

Managing the Crowd: Towards a Taxonomy of Crowdsourcing Processes. Presented at *the Seventeenth Americas Conference on Information Systems (AMCIS)*, Association for Information Systems, AIS Electronic Library (AISeL), Detroit, Michigan.

Georgakopoulou, Y., Sosoni, V., 2016. Crowdsourcing for the Creation of Parallel Translation Corpora: The Case of TraMOOC. Presented at: *Languages & The Media: 11th International Conference on Language Transfer and Audiovisual Media*. 03-04 November 2016. Berlin.

Gidron, B., 1985. Predictors of Retention and Turnover Among Service Volunteer Workers. *Journal of Social Service Research*. Volume 8, Issue 1, pp. 1-16.

Gile, D., 2004. Integrated Problem and Decision Reporting as a Translation Training Tool. *Journal of Specialised Translation*. Issue 2. [Online]. [Accessed on 26/11/17]. Available from http://www.jostrans.org/issue02/art_gile.php.

Görög, A., 2014. Quantifying and Benchmarking Quality: the TAUS Dynamic Quality Framework. *Tradumàtica*. No.12. pp. 443-454.

Görög, A., 2017. The 8 Most Used Standards and Metrics for Translation Quality Evaluation. [Online]. [Accessed on 26/11/17]. Available from http://blog.taus.net/the-8-most-used-standards-and-metrics-for-translation-quality-evaluation.

Hanneman, R.A., 2005. Introduction to Social Network Methods. Riverside, CA: University of California.

Hansen, G., 2009. A Classification of Errors in Translation and Revision. In: *CIUTI-Forum 2008: Enhancing Translation Quality: Ways, Means, Methods*. Peter Lang, pp. 313–326.

Harris, B., 1977. The Importance of Natural Translation. In: *Working Papers in Bilingualism*. No. 12. pp. 96-114.

Harris, B., Sherwood, B., 1978. Translating as an Innate Skill. In: Gerver D. and Sinaiko, H.W., eds., *Language Interpretation and Communication*. NATO Conferences Series. Springer, Boston, MA. Volume 6, pp.157-170.

Harris, B., 1980. How A Three-Year Old Translates. [Online]. [Accessed on

05/11/17]. Available from http://www.academia.edu/3846903/How_a_three-year-old_translates.

Hars, A., Ou, S., 2002. Working for Free? Motivations for Participating in Open-Source Projects. *International Journal of Electronic Commerce.* Volume 6, Issue 3, pp. 25–39.

Hennion, A., 2007. Those Things That Hold Us Together: Taste and Sociology. *Cultural Sociology*. Volume 1, Issue 1, pp. 97–114.

Hoffman, M.L., 1981. Is Altruism Part of Human Nature? *Journal of Personality and Social Psychology*. Volume 40, Issue 1, pp. 121–137.

House, J., 2001. Translation Quality Assessment: Linguistic Description versus Social Evaluation. *Meta*. Volume 46, Issue 2, pp 243-257.

Howe, J., 2006a. The Rise of Crowdsourcing. *Wired Magazine*.

Howe, J., 2006b. Crowdsourcing: A definition. *Crowdsourcing: Why the Power of the Crowd is driving the Future of Business*. [Online]. [Accessed on 05/11/17]. Available from http://crowdsourcing.typepad.com/cs/2006/06/.

Howe, J., 2008. Crowdsourcing: How the Power of the Crowd is Driving the Future of Business. Business Books. Great Britain.

Humphreys, A., Humphreys, J., 2013. Reading Difficulty Levels of Selected Articles in the Journal of Research in Music Education and Journal of Historical Research in Music Education. *Music Education Research International*. Volume 6. pp.15-25.

International Organisation for Standardisation, 2015. ISO 17100 Certified Translation Services. [Online]. [Accessed on 05/11/17]. Available from http://iso17100.com/.

Jenkins, H., 2006. Fans, Bloggers, and Gamers: Exploring Participatory Culture. NYU.

Jenkins, H., Deuze, M., 2008. Convergence Culture. *Convergence: The International Journal of Research into New Media Technologies.* Editorial 14(1), pp. 5–12.

Jenkins, H., Purushotma, R., Weigel, M., Clinton, K., Robison, A., 2009. Confronting the Challenges of Participatory Culture: Media Education for the 21st Century. The MIT Press.

Jiménez-Crespo, M.A., 2011. From Many One: Novel Approaches to Translation Quality in a Social Network Era. *Linguistic Antverpiensia. New Series – Themes Translation Studies*. No. 10. pp. 131-152.

Jiménez-Crespo, M.A., 2013a. Crowdsourcing, Corpus Use, and the Search for Translation Naturalness: A Comparable Corpus Study of Facebook and Non-Translated Social Networking Sites. *Translation and Interpreting Studies*. Volume 8, Issue 1, pp. 23–49.

Jiménez-Crespo, M.A., 2013b. Translation and Web Localisation. Routledge. New York & London.

Jiménez-Crespo, M.A., 2015. Testing Explicitation in Translation: Triangulating Corpus and Experimental Studies. *Across Languages and Cultures.* Volume 16, Issue 2, pp. 257–283.

Jiménez-Crespo, M.A., 2017. Crowdsourcing and Online Collaborative Translations: Expanding the Limits of Translation Studies, Benjamins Translation Library. John Benjamins Publishing Company, Amsterdam.

Kazman, R., Chen, H.M., 2009. The Metropolis Model a New Logic for Development of Crowdsourced Systems. *Communications of the ACM.* Volume 52, Issue 7, pp. 76–84.

Kelly, N., 2009. Freelance Translators Clash with LinkedIn over Crowdsourced Translation. *Common Sense Advisory.* [Online]. [Accessed on 15/11/17]. Available from http://www.commonsenseadvisory.com/Default.aspx?Contenttype=ArticleDetAD&tabID=63&Aid=591&moduleId=391.

Kelly, N., Ray, R., DePalma, D.A., 2011. From Crawling to Sprinting: Community Translation Goes Mainstream. *Linguistic Antverpiensia. New Series – Themes Translation Studies*. No. 10. pp.75-94.

Khalaf, A.S., Rashid, S.M., Jumingan, M.F., Othman, M.S., 2015. Problems in Amateur Subtitling of English Movies into Arabic. *Malaysian Journal of Languages and Linguistics*. Volume 3, Issue 1, pp. 38–55.

Kilgray., N.D. MemoQ 2014 R2 Help – Communication. [Online]. [Accessed on 11/03/17]. Available from http://kilgray.com/memoq/2014R2/help-full-en/index.html?memoq_help_title_page.html.

Kiraly, D., 2000. A Social Constructivist Approach to Translator Education. St. Jerome Publishing, Manchester.

Kosorukoff, A., 2011. Social Network Analysis: Theory and Applications. Passmore, D. L.

Kraut, R.E., 2017. The Development of L2 Reading Skills: A Case Study from an Eight-Week Intensive English Program Course. *Dialogues: An Interdisciplinary Journal of English Language Teaching and Research*. Volume 1, Issue 1, pp. 25-43.

Kunilovskaya, M., 2015. How far do we agree on the quality of translation? *English Studies at NBU*. Volume 1, Issue 1, pp.18–31.

Lakhani, K.R., Wolf, R., 2005. Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects. In: Feller. J., et al. eds., *Perspectives on Free Open Source Software*. MIT Press, Cambridge.

Language Realm, N.D. Words and Productivity for Translators. [Online]. [Accessed on 31/03/17]. Available from http://www.languagerealm.com/articles/word_productivity.php.

Leimeister, J.M., Huber, M., Bretschneider, U., Krcmar, H., 2009. Leveraging Crowdsourcing: Activation-Supporting Components for IT-Based Ideas Competition. *Journal of Management Information Systems*. Volume 26, Issue 1, pp. 197–224.

Lévy, P., 1999. Collective Intelligence: Mankind's Emerging World in Cyberspace. Basic Books, Cambridge.

Li, C., Bernoff, J., Fiorentino, R., Glass, S., 2007. Mapping Participation in Activities Forms the Foundation of a Social Strategy. *FORRESTER Social. Technographics*. [Online]. [Accessed on 31/03/16]. Available from https://www.forrester.com/report/Social+Technographics/-/E-RES42057.

Livingstone, S., 2004. The Challenge of Changing Audiences: or, what is the Audience Researcher to Do in the Age of the Internet? *European Journal of Communication.* Volume 19, Issue 1, pp. 75–86.

Lommel, A., Burchardt, A., Uszkoreit, H., 2014. Multidimensional Quality Metrics (MQM): A Framework for Declaring and Describing Translation

Quality Metrics. *Tradumàtica: tecnologies de la traducció*. No. 12, pp. 455–463.

Lommel, A., Ray, R., 2009. Crowdsourcing: The Crowd Wants to Help You Reach New Markets. Localisation Industry Standards Association.

Ma, J., 2017. The Benefits of Crowdsourcing Your Translation Project. [Online]. [Accessed on 15/02/18]. Available from http://blog.csoftintl.com/benefits-crowdsourcing-translation/.

Malakoff, M.E., Hakuta, K., 1991. Translation Skill and Metalinguistic Awareness in Bilinguals. In: Bialystok, E., eds., *Language Processing in Bilingual Children*. Cambridge University Press. pp. 141–165.

Malakoff, M.E., 1992. Translation Ability: A Natural Bilingual and Metalinguistic Skill. In: Harris, R.J., eds., *Advances in Psychology: Cognitive Processing in Bilinguals*. Volume. 83. pp.515-529.

Malone, T.W., Laubacher, R., Dellarocas, C., 2010. Harnessing Crowds: Mapping the Genome of Collective Intelligence. *MIT Slogan Research Paper*. No.4732-09.

Marin, A., Wellman, B., 2011. Social Network Analysis: An Introduction. In: Scott, J. and Carrington, P.J., eds., T*he SAGE Handbook of Social Network Analysis*. SAGE Publications Ltd. pp. 11–25.

Marshall, C., Rossman, G.B., 2016. Designing Qualitative Research. Sixth. edition. SAGE Publications.

Maslow, A.H., 1970. Motivation and Personality. Harper & Row.

Massey, G., Brändli, B., 2016. Collaborative Feedback Flows and How We Can Learn from Them: Investigating a Synergetic Learning Experience in Translator Education. In: Kiraly, D., eds., *Towards Authentic Experiential Learning in Translator Education*. Mainz University Press. pp. 177–199.

Matis, N., 2014. How to Manage Your Translation Projects. [Online]. [Accessed on 12/04/17]. Available from http://www.translation-project-management.com/book.

McAlpine, R., ND. From Plain English to Global English. [Online]. [Accessed on 08/07/17]. Available from http://www.angelfire.com/nd/nirmaldasan/journalismonline/fpetge.html.

McNab, R., Walters, K., 2016. Project Management Training at Sandberg Translation Partners Ltd. 04 May 2016. Whiteley, UK.

Melis, M.N., Hurtado, A. A., 2001. Assessment in Translation Studies: Research Needs. *Meta*. Volume 46, Issue 2, pp. 272-287.

Mesipuu, M., 2012. Translation Crowdsourcing and User-Translator Motivation at Facebook and Skype. *Translation Spaces*. Volume 1, Issue 1, pp. 33–53.

Mesipuu, M., 2011. Translation Crowdsourcing – an Insight into Hows and Whys (at the Example of Facebook and Skype). (Master Thesis). Tallinn University. Tallinn.

Meyer, B., Birte, P., Ortrun, K., 2010. Family Interpreters in Hospitals: Good Reasons for Bad Practice? *Mediazioni*. No. 10. pp297-324.

Meyer, D., 2011. Community Translation: Leveraging Communities and Crowds. Presented at *TAUS Roundtable on Collaborative Translation*. 10 October 2011. Santa Clara, USA.

Morera-Mesa, A., 2014. Crowdsourced Translation Practices: From the Process Flow Perspective. (Doctoral Thesis). Univeristy of Limerick.

Morningside Translations, 2016. Translation Quality Standards – What Do They Mean? [Online]. [Accessed on 26/11/17]. Available from https://www.morningtrans.com/translation-quality-standards-what-do-they-mean/.

Morschheuser, B., Hamari, J., Koivisto, J., 2016. Gamification in Crowdsourcing: A Review. In: *Proceedings of* the *49th Hawaii International Conference on System Sciences (HICSS)*, January 2016. pp. 4375–4384.

Mozilla, 2017. The History of MDN [Mozilla Developer Network]. [Online]. [Accessed on 26/11/17]. Available from https://developer.mozilla.org/en-US/docs/MDN_at_ten/History_of_MDN.

Munday, J., 2012. Introducing Translation Studies: Theories and Applications, 3rd edition. Routledge. New York & London.

Munro, R., 2010. Crowdsourced Translation for Emergency Response in Haiti: the Global Collaboration of Local Knowledge. In: *In Relief 2.0 in*

Haiti. [Online]. [Accessed on 26/11/17]. Available from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.224.6072.

Munro, R., Bethard, S., Kuperman, V., Lai, V.T., Melnick, R., Potts, C., Schnoebelen, T., Tily, H., 2010. Crowdsourcing and Language Studies: The New Generation of Linguistic Data. In: *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*. Association for Computational Linguistics, Stroudsburg, PA, USA, pp. 122–130.

Muthukumaraswamy, K., 2010. When the Media Meet Crowds of Wisdom. *Journalism Practice*. Volume 4, Issue 1, pp. 48–65.

National Translation Test and Appraisal Centre, 2017. CATTI Applicants in 2017. (2017 年全国翻译资格考试). [Online]. [Accessed 26.02.18]. Available from http://www.catti.net.cn/2017-11/06/content_750156.htm.

Neunzig, W., Tanqueiro, H., 2005. Teacher Feedback in Online Education for Trainee Translators. *Meta*. Volume 50, Issue 4, pp. 2005.

Nida, E. A., 1964. Towards a Science of Translating: With Special Reference to Principles and Procedures Involved in Bible Translating. Brill Archive.

Nord, C. 1997. Translating as a Purposeful Activity: Functionalist Approaches Explained. St Jerome. Manchester.

O'Brien, S., Schäler, R., 2010. Next Generation Translation and Localisation: Users are Taking Charge. *Translating and the Computer.* No. 32, pp.18-19.

O'Brien, S., 2012. Towards a Dynamic Quality Evaluation Model for Translation. *The Journal of Specialised Translation*. Issue 17. [Online]. [Accessed on 15/02/18]. Available from: http://www.jostrans.org/issue17/art_obrien.php.

O'Hagan, M., 2008. Fan Translation Networks: An Accidental Translator Training Environment? In: Kearns, J., eds., *Translator and Interpreter Training: Issues, Methods and Debates*. A&C Black. pp:159-183.

O'Hagan, M., 2009. Evolution of User-generated Translation: Fansubs, Translation Hacking and Crowdsourcing. *The Journal of*

Internationalisation and Localisation. Volume. 1, pp. 94–121.

O'Hagan, M., 2011. Community Translation: Translation as a Social Activity and Its Possible Consequences in the Advent of Web 2.0 and Beyond. *Linguistica Antverpiensia, New Series – Themes in Translation Studies*. No. 10, pp. 11-23.

O'Hagan, M., 2013. The Impact of New Technologies on Translation Studies. In: Millán, C. and Bartrina, F., eds., *Routledge Handbook of Translation Studies*. Routledge. New York & London. pp. 503–518.

O'Hagan, M., Mangiron, C., 2013. Game Localisation: Translating for the Global Digital Entertainment Industry. John Benjamins Publishing Company, Amsterdam.

Olohan, M., 2012. Volunteer Translation and Altruism in the Context of a Nineteenth-Century Scientific Journal. *The Translator.* Volume 18, Issue 2, pp. 193–215.

Olohan, M., 2014. Why do you Translate? Motivation to Volunteer and TED Translation. *Translation Studies*. Volume 7, Issue 1, pp. 17–33.

Orrego-Carmona, D., 2012. Internal Structures and Workflows in Collaborative Subtitling. In: Antonini, R. and Bucaria, C., eds., *Non-Professional Interpreting and Translation*. Peter Lang, Frankurt. Vol 8, pp. 231-256.

Orrego-Carmona, D., 2014. Using Non-Professional Subtitling Platforms for Translator Training. *Rivista internazionale di tecnica della traduzione. EUT* Edizioni Università di Trieste. No. 15. pp.129-144.

Orrego-Carmona, D., 2015. The Reception of (Non)Professional Subtitling. (Doctoral Thesis). UNIVERSITAT ROVIRA I VIRGILI.

Ozinga, J.R., 1999. Altruism. Greenwood Publishing Group.

Packham, S., 2016. Crowdsourcing a Text Corpus for a Low Resource Language. (Master Thesis). University of Cape Town.

Pérez, E., Carreira, O., 2011. Evaluación del modelo de crowdsourcing aplicado a la traducción de contenidos en redes sociales: Facebook. In Encinas, E.C. et al. eds, *La Traductología Actual: Nueva Vías de Investigación en la Disciplina*. Granada: Comares. pp. 99–118.

Pérez-González, L., Susam-Saraeva, Ş., 2012. Non-professionals

Translating and Interpreting. *The Translator.* Volume 18, Issue 2. pp. 149–165.

Perrino, S., 2009. User-Generated Translation: The Future of Translation in a Web 2.0 Environment. *The Journal of Specialised Translation*. Issue 12. pp. 55–78.

Post, M., Callison-Burch, C., Osborne, M., 2012. Constructing Parallel Corpora for Six Indian Languages via Crowdsourcing. In: *Proceedings of the Seventh Workshop on Statistical Machine Translation*, Association for Computational Linguistics, Stroudsburg, PA, USA. pp. 401–409.

Prioux, R., Michel, R., 2007. Economie de la révision dans une organisation internationale : le cas de l'OCDE. *The Journal of Specialised Translation*. Issue 8. pp. 21–41.

Proz, 2009. Professional Translators Against Crowdsourcing and Other Unethical Business Practices. *Translation – Art &Business – Business Issues*. [Online]. [Accessed on 13/12/27]. Available from: https://www.proz.com/forum/business_issues/153668-professional_translators_against_crowdsourcing_and_other_unethical_business_practices.html.

Pym, A., 2011. Translation Research Terms: A Tentative Glossary for Moments of Perplexity and Dispute. In: Pym, A., eds., *Translation Research Projects*. Tarragona: Intercultural Studies Group, Volume 3. pp. 75–99.

Rajala, R., Westerlund, M., Vuori, M., Hares, J.-P., 2013. From Idea Crowdsourcing to Managing User Knowledge. *Technology Innovation Management Review*. December 2013. pp. 23–31.

Ray, R., Kelly, N., 2011. Crowdsourced Translation: Best Practices for Implementation. *Common Sense Advisory*. February 2011.

Raymond, E.S., 2001. The Cathedral & the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary. 1st edition. O'Reilly Media, Beijing, Cambridge, Mass.

Reiss, K., 1989. Text Type, Translation Types and Translation Assessment. Translated by Chesterman A., In: Chesterman, A., eds., *Readings in*

Translation Theory. Oy Finn Lectura Ab. Helsinki. pp.105-115

Rouse, A., 2010. A Preliminary Taxonomy of Crowdsourcing. In *ACIS 2010 Proceedings*. 76. [Online]. [Accessed on 12/11/17]. Available from http://aisel.aisnet.org/acis2010/76/.

Ryan, R.M., Deci, E.L. 2000. Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*. Volume 25, Issue 1, pp. 54-67.

Sajan. 2013. Translation Crowdsourcing: Weighing the Pros and Cons. [Online]. [Accessed on 15/02/18]. Available from https://www.sajan.com/translation-crowdsourcing-weighing-the-pros-and-cons/.

Schäffner, C., 1998. From 'Good' to 'Functionally Appropriate': Assessing Translation Quality. In: Schäffner, C., eds., *Translation and Quality*. Clevedon: Multilingual. pp. 1-5.

Schwimmer, M., 2017 Beyond Theory and Practice: Towards an Ethics of Translation. *Ethics and Education*. Volume 12, Issue 1, pp.51-61.

Scott, J., Carrington, P.J. (Eds.), 2011. The SAGE Handbook of Social Network Analysis, 1st edition. SAGE Publications Ltd. London, Thousand Oaks, CA.

Secară, A., 2005. Translation evaluation–a state of the art survey. In: *Proceedings of the ECoLoRe/MeLLANGE Workshop*. Leeds. pp. 39–44.

Sheldon, K.M., Elliot, A.J., 1998. Not all Personal Goals are Personal: Comparing Autonomous and Controlled Reasons for Goals as Predictors of Effort and Attainment. *Personality and Social Psychology Bulletin*. Volume 24, Issue 5, pp. 546–557.

Shreve, G.M., 1997. Cognition and the Evolution of Translation Competence. In: Danks, J.H. et al. eds., *Cognitive Processes in Translation and Interpretation.* Thousand Oaks CA: Sage Publications. pp.120-136.

Shreve, G.M., 2006. The Deliberate Practice: Translation and Expertise. *Journal of Translation Studies*. Volume 9, Issue 1, pp.27-42.

Shye, S., 2009. The Motivation to Volunteer: A Systemic Quality of Life Theory. *Social Indicators Research*. Volume 98, Issue 12. pp. 183–

200.

Soller, A., 2001. Supporting social interaction in an intelligent collaborative learning system. *International Journal of Artificial Intelligence in Education.* 12, pp. 40–62.

State Administration for Quality Supervision and Inspection, 2005, 2008. Specification for Translation Service—Part 1: Translation (中国语言服务行业规范). [Online]. [Accessed on 26/11/17]. Available from http://tac-online.org.cn/ch/tran/2013-04/19/content_5890461.htm.

Stewart, O., Lubensky, D., Huerta, J.M., 2010. Crowdsourcing Participation Inequality: A SCOUT Model for the Enterprise Domain. In: *Proceedings of the ACM SIGKDD Workshop on Human Computation*, ACM Digital Lirary, New York, NY, USA, pp. 30–33.

STP, N.D. Revision | Translation Services Procedure. [Online]. [Accessed on 31/03/17]. Available from http://stptrans.com/processes-standards-qc/how-we-work/revision/.

Strauss, A., Corbin, J.M., 1998. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. SAGE Publications.

TAUS, 2015. Translation Productivity Revisited. [Online]. [Accessed on 31/03/17]. Available from https://www.taus.net/blog/translation-productivity-revisited.

Teasley, S.D., Roschelle, J., 1993. Constructing a Joint Problem Space: The Computer as a Tool for Sharing Knowledge. In O'Malley, C., eds., *Computer-Supported Collaborative Learning*. Springer-Verlag, New York. pp. 229–258.

Tran, Y., Yonatany, M., Mahnke, V., 2016. Crowdsourced Translation for Rapid Internationalisation in Cyberspace: A Learning Perspective. *International Business Review*. Volume 25, Issue 2. pp. 484–494.

Trompette, P., Chanal, V., Pelissier, C., 2008. Crowdsourcing as a Way to Access External Knowledge for Innovation. Presented at *the 24th EGOS Colloquium*. July 2008. Amsterdam, France.

Tsang, F., 2008. Crowdsourcing at Adobe. *TAUS*. [Online]. [Accessed on 13/12/17]. Available from https://www.taus.net/think-tank/news/news-and-alerts/crowdsourcing-at-adobe.

Tsvetkov, N., Tsvetkov, V., 2011. Effective Communication in Translation and Localisation Project Management. In: Dunne, K.J. and Dunne, E.S., eds., *Translation and Localization Project Management: The Art of the Possible*. John Benjamins Publishing Company. Amsterdam. pp. 189–210.

Vermeer, H.J., 2004. Skopos and Commission in Translational Action. In Venuti, L., eds., *The Translation Studies Reader*. Routledge. New York & London. pp. 221-232.

VideoLocalise, 2017. Video Localisation: Driving Down Costs, Enabling Scale. White Paper, *Multilingual*, September 2017. pp 32-33.

Vlachopoulos, G., 2009. Translation, Quality and Service at the European Commission. In: *CIUTI-Forum 2008: Enhancing Translation Quality: Ways, Means, Methods*. Peter Lang.

Vukovic, M., 2009. Crowdsourcing for Enterprises. In: *Services-I, 2009 Congress on Services*. IEEE, pp. 686–692.

Waddington, C., 2001. Should Translations be Assessed Holistically or through Error Analysis? *Journal of Language and Communication*. Hermes. No. 26. pp. 15–37.

Waldron, J., 2013. User-generated Content, YouTube and Participatory Culture on the Web: Music Learning and Teaching in Two Contrasting Online Communities. *Music Education Research*. Volume 15, Issue 3, pp. 257–274.

Whyatt, B., 2012. Translation as a Human Skill: From Predisposition to Expertise. Publikacja dofinansowana przez. POZNAŃ.

Whyatt, B., 2017. Chapter 3. We are All translators: Investigating the Human Ability to Translate from a Developmental Perspective. In: Antonini, R., et al. eds., *Non-Professional Interpreting and Translation: State of the Art and Future of an Emerging Field of Research*. John Benjamins Publishing Company, Amsterdam/Philadelphia.

Williams, J., Chesterman, A., 2014. The Map: A Beginner's Guide to Doing Research in Translation Studies. Routledge. New York & London.

Williams, M., 2001. The Application of Argumentation Theory to Translation Quality Assessment. *Meta*. Volume 46, Issue 2, pp 326-344.

Williams, M., 2004. Translation Quality Assessment: An Argumentation-Centred Approach. University of Ottawa Press. Ottawa.

Xu, R., Sharoff, S., 2014. Evaluating Term Extraction Methods for Interpreters. Presented at: *the 4th International Workshop on Computational Terminology (Computerm)*. Aclweb. Dublin, Ireland.

Yahaya, F., 2008. Managing complex translation projects through virtual spaces: a case study. *ASLIB Translating and the Computer*. Volume 30. pp. 27–28.

Yang, X., 2010. The 39 Days Since Yeeyan Left (译言走后的 39 天). *Southern People Weekly Magazine*. [Online]. [Accessed on 14/09/16]. Available from http://www.infzm.com/content/40739.

Yeeyan, 2008. Earthquake Safety Manual (地震安全手册). Earthquake Publishing. [Online]. [Accessed on 14/09/16]. Available from https://book.douban.com/subject/1729172/.

Yeeyan, 2016. The Rectification Notice on the Recent Changes of Webpage and User System (关于译言网站近期版面和用户系统调整的公告). [Online]. [Accessed on 14/09/16] Available from http://article.yeeyan.org/view/137424/496528.

Yip, P. C., Don, R., 2004. Chinese: A Comprehensive Grammar. 1st Edition. Routledge Taylor&Francis Group, London and New York.

Yuan, Y., Sharoff, S., Babych, B., 2016. MoBiL: A Hybrid Feature Set for Automatic Human Translation Quality Assessment – Semantic Scholar. In: *the LREC 2016 Proceedings*. European Language Resources Association, pp. 3663–3670.

Yunker, J., 2014. Make your Product 'Translation Worthy', and the World Will Follow. [Online]. [Accessed on 25/01/18]. Available from https://gigaom.com/2014/01/12/make-your-product-translation-worthy-and-the-world-will-follow/.

Zaidan, O.F., Callison-Burch, C., 2011. Crowdsourcing Translation: Professional Quality from Non-Professionals. In: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics. pp. 1220–1229.

Zhang, W., Mao, C., 2013. Fan Activism Sustained and Challenged: Participatory Culture in Chinese Online Translation Communities. *Chinese Journal of Communication.* Volume 6, Issue 1. pp. 45–61.

Zhao, Y., Zhu, Q., 2012. Evaluation on Crowdsourcing Research: Current Status and Future Direction. *Information Systems Frontiers*. Volume 16, Issue 3. pp. 417–434.

Zheng, H., Li, D., Hou, W., 2011. Task Design, Motivation, and Participation in Crowdsourcing Contests. *International Journal of Electronic Commerce*. Volume 15, Issue 4. pp. 57–88.

Appendix 1.

Questionnaire Survey for Members in *Yeeyan* Community

Hi, *Yeeyan* members, you are invited to participate in a study of crowdsourcing translation. It is used to explore the motivations that drive people to volunteer to translate in the online environment. The survey also includes questions regarding the preference and productivity of translators in *Yeeyan*, as well as translation strategy and quality control methods. This survey will take approximately 15 minutes and all data will be completely anonymised and securely protected. The open questions can be answered in either *English* or *Chinese*. Please be aware that, by completing the questionnaire survey, you give the ***consent*** for using data in this research, including publication of direct quotations of the comments you have made in the survey. Thanks so much for your contribution!

Researcher: Jun Yang

School of Languages, Cultures and Societies, University of Leeds

If you have any questions, please feel free to contact.

Email: ml11j3y@leeds.ac.uk

1. How old are you? (Single Choice)
   - ○ below 20
   - ○ 20-25
   - ○ 26-30
   - ○ above 30
   - ○ I prefer not to tell.

2. Your education background: (Single Choice)
   - ○ High School
   - ○ Bachelor
   - ○ Master
   - ○ PhD

3. Your English competence: (Multiple Choice)
   - □ CET 4: Pass
   - □ CET 4: Merit
   - □ CET 6: Pass
   - □ CET 6: Merit
   - □ TEM 4: Pass

□ TEM 4: Merit
□ TEM 8: Pass
□ TEM 8: Merit
□ IELTS: 6-6.5
□ IELTS: 7 or above
□ TOEFL: 90-100
□ TOEFL: 100 or above
□ None
□ Other _____

4. Did you major in language-related studies? (Single Choice)
○ Yes
○ No

5. Have you received any translation-related training? (Single Choice)
○ Yes
○ No

6. If ‘yes’ in question 5, what kind of translation-related training did you receive? Please describe it.

_____

7. Your current occupation: (Single Choice)
○ Student
○ Translator
○ Language-related work
○ Other _____

8. In which year did you become a member of *Yeeyan*?

_____

9. What is your language combination? (Multiple Choice)
□ English-Chinese
□ Japanese-Chinese
□ German-Chinese
□ Korean-Chinese
□ Russian-Chinese
□ French-Chinese
□ Other _____

10. What kind of roles have you fulfilled in *Yeeyan*? (Multiple Choice)

□ Active reader. I like reading articles in *Yeeyan*.
□ Mentor. I advise others on translation strategies or tools that they could use, and I answer their questions regarding translation practices.
□ Commentator. I like commenting on the article or others' translation.
□ Project manager. I like organising and managing a team project.
□ Reviser. I like revising members' work.
□ Promoter. I usually share translations from *Yeeyan* to other sites/sources.
□ Source text provider. I like recommending source texts to others.
□ Translator. I like transferring a text from one language to another.
□ Other _____

11. Which of the following are your top 5 reasons for participating in order of their importance, starting with the most important one? (Please select and rank them in the order of importance.)

_____To improve my language abilities.
_____To gain translation experience.
_____To make foreign information available to Chinese readers.
_____To help support the initiative of *Yeeyan*.
_____For fun.
_____To enhance my reputation as a translator.
_____To network with others and make more friends.
_____I need the information to be available in the target language.
_____To gain recognition within the community.
_____To get an early look at translation environment.
_____To attract translation clients.
_____Because the project is intellectually stimulating.
_____To earn money.
_____Because I am interested in the topic of the source text.
_____For the sense of self-achievement.
_____To voice my opinions on certain topics through my translation.
_____To gain the opportunity to have my translation published.
_____To earn honorable badges in *Yeeyan*.

12. Please write down any other reason of why you are participating if it is not listed in the previous question.

_____

13. In your opinion, what are the possible reasons that others participate in *Yeeyan*? You can either select the possible reasons or make your own comments. (Multiple Choice)

□ To improve language abilities.
□ To gain translation experience.
□ To make foreign information available to Chinese readers.
□ To help support the initiative of *Yeeyan*.
□ For fun.
□ To enhance one' reputation as a translator.
□ To network with others and make more friends.
□ The need to make the information available in the target language.
□ To gain recognition within the community.
□ To get an early look at translation environment.
□ To attract translation clients.
□ Because the project is intellectually stimulating.
□ To earn money.
□ Because the topic of the source text is interesting.
□ For the sense of self-achievement.
□ To voice opinions on certain topics through one's translation.
□ To gain the opportunity to have one's translation published.
□ To earn honorable badges in *Yeeyan*.
□ Other _____

14. In your experience, what may de-motivate your participation? For example, what are the reasons that may stop you from actively participating or lead you to quit in the middle of a project? (Multiple Choice)

□ The source text is too difficult.
□ There are too many rules to follow.
□ It is too hard to communicate with group members.
□ The process takes too much time.
□ The group members' performance is not good enough.
□ Other _____

15. What kind of articles/texts do you usually translate? (Multiple Choice)

□ Business
□ Literature
□ Media
□ Health
□ Popular Science
□ Politics
□ Technology
□ Other _____

16. How do you choose the articles that you want to translate? (Multiple Choice)

□ Based on the topic of the article.
□ Based on the source of the article.
□ Based on the members that I will collaborate with.
□ Other _____

17. How many translations (translation for articles) have produced in *Yeeyan*? (Single Choice)

○ None
○ Fewer than 10 translations
○ 10-30
○ 30-60
○ More than 60

18. How many hours per week on average do you spend in translation in *Yeeyan*?

_____

19. How do you think of the group collaborative translation in *Yeeyan*? (Single Choice)

○ I will never consider it. I prefer to translate all by myself.
○ I sometimes participate in it, if the chance arises.
○ I actively seek to participate. I like collaborative translation projects with discussion and feedback.
○ This is the only mode I would choose for all my translation work.

20. Could you please name a few *Yeeyan* 'interest groups' (译言小组) that you visit often or you like the most?

_____

21. Do you know or are you following any 'famous' translators in the *Yeeyan* community? If so, please name them.

_____

22. Have you earned any 'honourable badges' (译言勋章) from *Yeeyan*? (Single Choice)

○ Yes
○ No

23. Have you participated in the *Yeeyan Gutenberg Project* (译言古登堡计划)? (Single Choice)

○ Yes
○ No

24. If 'Yes' in question 23, what kind of role did you play? (Multiple Choice)

□ Project manager.
□ Reviser.
□ Translator.

25. During translation, when you come across a problem, such as you cannot understand the source text, you are not sure of the proper terminology to use, etc., what do you do to solve the problem? (Multiple Choice)

□ Consult subject-matter professionals.
□ Do further research on the subject by yourself.
□ Communicate with other translators/team members.
□ Ignore the problem/Not do anything about it.
□ Other _____

26. Have you used machine translation (机器翻译) (e.g. Google Translate, Baidu Translate, Youdao Translate) for your task in *Yeeyan*? (Single Choice)

○ Yes, I have used machine translation and I have left the translation as the machine suggested it.
○ Yes, I have used machine translation and I post-edited it to improve it.
○ Yes, I have used machine translation and I use it as a reference to lookup vocabulary, specialised terminology and expressions.
○ No, I have never used machine translation.

27. When you translate, what is your strategy? (Single Choice)

○ I try to stay as close as possible to the source text (尽可能与原文保持一致).
○ I usually translate and edit as I wish (按照个人想法进行编译).
○ I translate according to the project brief (根据项目要求进行翻译).
○ Other _____

28. When you have finished the first draft of your translation, will you self-revise (自我校对) your work? (Single Choice)

○ No, I won't.
○ Yes, I will read through my work again and make corrections immediately after my translation is finished.
○ Yes, I will leave it for a few days and read again to see if there is anything that needs to be improved.

29. If you revise your own work, what areas do you find yourself changing most often? (Multiple Choice)

□ General vocabulary (用词方面)

□ Specialised terminology (专业术语)

□ Grammar (语法)

□ Punctuation (标点)

□ Misspelled character (错别字)

□ Style and register (翻译风格和语域)

□ Inconsistency within the text (译文中的一致性)

□ Other _____
□ I never revise my own work.

30. After your translation is published, will you revisit your work to check other people's comments on it? (Single Choice)
○ Yes
○ No

31. In a collaborative project, will you go through other members' comments or revisions on your work? And do you find them helpful? (Single Choice)
○ No, when my translation is done, my job is done.
○ Yes, I find other people's suggestions very helpful and I can learn from them.
○ I have never participated in a collaborative project.

32. How do you revise others' translation? (Single Choice)

○ Spot check (节选校对). (e.g. I randomly select part of the translation to revise.)

○ Full check (全文校对). (e.g. I revise the translation from the beginning to the end.)
○ I have no experience in revising other's work.

33. When you revise other people’ s work, what is your procedure? (Single Choice)

○ Monolingual review (单语校对，只读译文), I only read the translation to check if there is any awkwardness or difficulties in understanding.
○ Comparative revision (对照原文，对比校对), I usually compare the translation with the source text to check if there is any mis-translations or other content-related problems.
○ Both. I will do a monolingual review first, and if there are too many problems, I will consult the source text and do a comparative revision.
○ Both. I will do a comparative revision first to check the content, and then a monolingual review to check the language.
○ Other _____
○ I have no experience in revising other's work.

34. In the past collaborative projects that you have participated in, have you ever been asked to do a sample test (试译) before you joined the group? (Single Choice)

○ Yes. I did a sample test in order to participate in the project.
○ No. I just signed in and started my work.
○ I have never participated in a collaborative project.

35. In the past collaborative projects that you have participated in, what are the possible methods that you used to ensure consistency (保障译文的一致性)? (Multiple Choice)

□ Using a glossary list.
□ Frequent discussion with team members.
□ A subject-matter professional/project manager to do the final review.
□ Other _____
□ I have never participated in a collaborative project.

36. In the past collaborative projects that you have participated in, is there a lessons-learned session (总结学习环节) organised by the team leader or project manager after the translation is finished? (Single Choice)

○ Yes
○ No
○ I have never participated in a collaborative project.

37. When you are involved in a collaborative project, will you look through the brief or requirements that the organiser/project manager produced before you start translating? (Single Choice)

○ Yes, I look through the brief and I always follow the requirements.
○ No, I only do my part of the task.
○ Yes, I look through the brief and I will communicate with the project manager if there is anything that I do not understand, or I think is negotiable.
○ I have never participated in a collaborative project.

38. If something unexpected occurs that makes you quit in the middle of the project or may cause the delay of your work, will you give the team members/project manager a notice about this? (Single Choice)

○ Yes, I will.
○ No, I will not.
○ It depends (please explain). _____
○ I have never participated in a collaborative project.

39. Have you ever disappeared/quitted in a project before completing your task? (Single Choice)

○ Yes.
○ No.
○ I have never participated in a collaborative project.

40. Do you like communicating with team members in a collaborative project? (Single Choice)

○ Yes, I like communication very much. It helps improve the quality of my work.
○ Yes, I like communication very much. It helps me bonding with other members.
○ No, I do not like communicating with others and I prefer to focus on my own task.
○ I have never participated in a collaborative project.

41. What are the possible reasons that make you follow the rules in the project and actively collaborate with others?

□ When joining the project, I have made a commitment and I need to fulfill that.
□ I believe that rules and collaboration can produce better quality.
□ I see the work as a mock project in the professional world and I want to learn from it.
□ Other _____

□ I have never participated in a group collaborative translation project.

42. Please think about your experience in *Yeeyan*, what have you gained?

_____

43. What do you consider to be your contribution to the *Yeeyan* community?

_____

44. According to your experience in *Yeeyan*, is there anything that you think can be improved? You may write down the expectations for yourself or for *Yeeyan*.

_____

Appendix 2.

Specifications of the collaborative translation projects observed in the study

Health Project I	
External Features of Source Text	
Author	Bonnie Goldman
Title	The First Man to Be Cured of AIDS: An Update on the Amazing Story – This Month in HIV
Language variety	English (US), formal journalism and academic writing.
Genre or text type	Popular science report of specialised article in the health domain. Informative text type
Length	10,607 words
Publication outlet	The Body (The Complete HIV/AIDS Resource)
Readership	Generally lay, educated, non-expert
Place of publication	Online http://www.thebody.com/content/art53624.html?getPage=1
External Features of Target Text	
Language variety	Chinese (PRC)
Length	Full text translation
Translator's notes (footnotes or endnotes)	Acceptable
Readers' knowledge of subject	Generally lay, educated, non-expert
Publication outlet	*Yeeyan*
Collaboration platform	http://sync.in/sFIooyhScG ; http://sync.in/dK9Vno0MJi
Place of publication	Online http://article.*Yeeyan*.org/view/231604/328757
Health Project II	

External Features of Source Text	
Author	Siddhartha Mukherjee
Title	Do Cellphones Cause Brain Cancer?
Language variety	English (US), formal journalism and academic writing.
Genre or text type	Popular science report of specialised article in the health domain. Informative text type
Length	5,389 words
Publication outlet	The New York Times
Readership	General lay, educated, non-expert
Place of publication	Online http://www.nytimes.com/2011/04/17/magazine/mag-17cellphones-t.html?_r=2&ref=technology
External Features of Target Text	
Language variety	Chinese (PRC)
Length	Full text translation
Translator's notes (footnotes or endnotes)	Acceptable
Readers' knowledge of subject	Generally lay, educated, non-expert
Publication outlet	*Yeeyan*
Collaboration platform	http://sync.in/oITetldYuU
Place of publication	Online http://article.*Yeeyan*.org/view/147927/190347
Business Project I	
External Features of Source Text	
Author	Alex Nussbaum, David Voreacos, and Greg Farrell
Title	Johnson & Johnson's Quality Catastrophe
Language variety	English (US), formal journalism and academic writing.

Genre or text type	Popular science report of specialised article in the business domain. Informative text type
Length	6,080 words
Publication outlet	Bloomberg Businessweek
Readership	Generally lay, educated, non-expert
Place of publication	Online http://article.*Yeeyan*.org/view/147927/188898#top
External Features of Target Text	
Language variety	Chinese (PRC)
Length	Full text translation
Translator's notes (footnotes or endnotes)	Acceptable
Readers' knowledge of subject	Generally lay, educated, non-expert
Publication outlet	*Yeeyan*
Collaboration platform	http://sync.in/2c4J1fhFXF
Place of publication	Online http://article.*Yeeyan*.org/view/147927/188898#top
Business Project II	
External Features of Source Text	
Author	Mark Clothier
Title	Warren Buffett's Ride on the Rails Is Paying Off
Language variety	English (US), formal journalism.
Genre or text type	Popular science report of specialised article in the business domain. Informative text type
Length	962 words
Publication outlet	Bloomberg Businessweek
Readership	Generally lay, educated, non-expert

Place of publication	Online https://www.bloomberg.com/news/articles/2011-05-12/warren-buffetts-ride-on-the-rails-is-paying-off
External Features of Target Text	
Language variety	Chinese (PRC)
Length	Full text translation
Translator's notes (footnotes or endnotes)	Acceptable
Readers' knowledge of subject	Generally lay, educated, non-expert
Publication outlet	*Yeeyan*
Collaboration platform	http://sync.in/VCF*GP*8nF9r
Place of publication	Online http://article.*Yeeyan*.org/view/147927/196901#top
Gutenberg Project	
External Features of Source Text	
Author	Louisa May Alcott
Title	Hospital Sketches
Language variety	English (US), literary text.
Genre or text type	Classic aesthetic literature. Expressive text type.
Length	28,886 words
Publication outlet	Public domain – Project Gutenberg
Readership	Generally lay
Place of publication	Online, e-book http://www.gutenberg.org/ebooks/3837
External features of Target Text	
Language variety	Chinese (PRC)
Length	Full text translation
Translator's notes	Acceptable

(footnotes or endnotes)	
Readers' knowledge of subject	Generally lay
Publication outlet	*Yeeyan* and *Douban*
Collaboration platform	None
Place of publication	Online, e-book https://read.douban.com/ebook/23313484/

Appendix 3

Questions in the follow-up interviews

Semi-structured interviews with participants in the *Gutenberg Project*

1. Why do you take part in the project?
2. Is there any standard that you have for translation and revision? What does the standard include? What do you pay attention to in each procedure?
3. How do you revise your work or your team member's work? Monolingual review or comparative revision?
4. How do you do information mining in the project? For instance, background research, terminology search, etc.?
5. Do you find a big difference between the draft translation, self-revised version and the final peer-revised version?
6. Do you like communicating with your team members? How useful is it?
7. How do you evaluate your performance in this project? Can you comment on your team members' performance?
8. What have you gained from this project? Can you think of any improvements that you realised after the project?
9. Are there any new things that you want to attempt in your next project or anything that you think can be improved?