Qualitative Research Methods in Second Language Learning:Review and Evaluation

Clare F. Kaneko, Keiwa College

The first time you attempt a qualitative research project it is difficult to assess the effectiveness of the methods you use until the research project itself is complete. My intention in writing this paper is to provide reflections and insights into my thinking about the choice of research methods for my Masters dissertation, as well thoughts for future projects using these same methods. A combination of qualitative methods was used for my research project, including think aloud protocol, interviews, and stimulated recall. By reflecting critically on the implementation of these methods, I was able to identify various areas that could be improved on in future projects. Using qualitative research methods and adjusting them specifically for the project to be carried out, researchers can gather more useful data, but the resulting increase in the amount of data that needs to be analyzed can cause a project to consume more time that was originally envisaged. It is hoped that by reading this paper novice researchers may gain some insight in how to approach their own research using qualitative methods.

初めて質的研究のプロジェクトを行う場合、研究自体が完了するまでは、そのプロジェクトの手法の有効性を評価することは困難である。本稿の目的は、筆者の修士論文の研究手法の選択について省察と洞察を与え、また、同じ手法を用いる今後のプロジェクトに示唆を与えるものである。本研究プロジェクトでは、定性的方法の組み合わせ（考え話す記録法、インタビュー法、刺激想起法）を使用した。これらの方法を批判的に省察することによって、今後のプロジェクトで改善可能な部分を明らかにすることができた。定性的な研究手法を用い、それを特定のプロジェクトに合わせて調整することで、研究者はより有用なデータを収集することができるが、プロジェクトの時間を浪費させるのはデータの量である。本稿は初めてリサーチを行う研究者に、質的手法を用いた自身の研究のアプローチに関する洞察を与えるであろう。

Keywords

reflective, qualitative, effectiveness, research methods

キーワード

省察的、定性的、有効性、調査方法

Attempting a research project for the first time can be quite daunting. I know that was certainly true for me. Where do you start, and how do you find the information you need? This paper provides firsthand account of the methods and techniques chosen for a first–time qualitative research project. The structure of this paper follows the chronological order in which the project was completed: pre–project decisions, project design and implementation, and post–project reflection. This last section examines the effectiveness of the methods and techniques in obtaining the necessary data in relation to the research questions and notes how these could have been better applied with hindsight.

My first step in the project was to refer to existing reference books on qualitative research methods. There are many general or genre–specific "how to books" on qualitative research methods in social work (Padgett, 2017), psychology (Willig, 2013), or applied linguistics (Dörnyei, 2007). I do not attempt to add to this collection of "how to books", but instead provide a reflection of my own learning about these methods as I applied them in my own research. My reflection on the project used was undertaken 6 months after the research had been completed giving me time to process how well the research had been carried out. As a first–time student–researcher, I discovered that the literature I read focused on description and what the qualitative method could achieve, as well as the pros and cons of use, rather than practical implementation. My aim, therefore, is to describe the methods that I chose and

explain how they were implemented in a qualitative research project. The reflection on their implementation allowed for greater learning and considerations for future projects. I knew that my research project was not flawless, and my desire to know where I could have improved led me to write this paper. The shortfalls discussed when using these methods are not necessarily due to the method itself but, rather, my interpretation and application of them.

Pre-Project Decisions

The project I reflect on in this paper was my dissertation for the Birmingham Masters of Arts Course in TEFL/ TESL in 2017. Masters students are expected to write a 15,000-word research paper under the guidance of a supervisor. The research project was to be completed within six months of initial contact with the supervisor. I started my masters after working for 15 years in English education in Japan, including six months at the tertiary level before undertaking my dissertation. As an avid user of technology, in all of my classes I encourage students to use whatever tools are available to them to enhance their learning. The choice of topic for my dissertation arose from watching how the low proficiency students at a small private Japanese university used their smart phones when creating scripts for conversations, rather than rely on their own ability. The less proficient students appeared to rely on their phones more than their textbooks. Unfortunately, when checking the English that my students wrote, I found simple mistakes that even lower proficiency students should not make. In addition, there were more complex errors that often rendered the conversation incomprehensible. As I was unsure how the mistakes were made and use of smart phones is ever increasing, I felt this was an area of personal interest that I could do research into. How can advancing technology not improve how we do things, or in fact make them worse?

   I firstly reviewed what software the students were using for writing their dialogues and compared their favorite translation apps with Google Translate. This enabled the students to see firsthand how different software produces different results. Students decided that Google Translate was the most accurate app and they consequently used it to help them with translation. This alone did not solve all the translation problems, and a new concern that students may use their phones in other ways to find the information they were looking for via YouTube or online searches helped me further focus my research area on how students used smart phones for producing written language.

   At this point, I sent my dissertation proposal to Birmingham University and subsequently a supervisor was assigned to work with me. Our first contact focused on why students used their smart phones for creating conversations. I had not thought much of the wording of the research question or with what methods I would attempt to answer it. I did know that I wanted to answer the *why* question: Why did students choose to use their smart phones for a particular question, and in what ways other than translation were they using their phones for? We agreed that the best way to gain access to the thoughts of the participants was to use a qualitative approach. My supervisor suggested I look further into Second Language Acquisition research methods and provided suggestions of books (Dörnyei, 2007; Ellis & Barkhiuzen, 2005) to aid me in the design of my project. Indeed, my supervisor was an important influence on me throughout the project, so it is worth saying a little more about this relationship.

Supervisor—Supervisee Relationship

Throughout the design of the research project I was in contact with my supervisor. Not only did my supervisor advise me on the general steps and processes of research and writing, he also suggested ideas that were not outlined in the "how to books" I had read. Working through the project, my supervisor suggested various research method books that would help

me to understand the methods available. In turn, I attempted to learn more about each method and how to implement it. As the method of data collection is important in qualitative research, being able to discuss with someone who has used various methods in their own projects was as valuable as reading academic papers and books. We communicated by Skype and email. My supervisor consistently provided suggestions rather than directives, ensuring that I looked for information on how to progress on my own.

Project Design and Implementation

To answer my original *why* question, it was necessary for me to create more specific and detailed questions designed to help me describe what students do, when and how they do it:

1. *To what extent do low-level Japanese tertiary English language learning students rely on their smart phones in written language production?*
2. *When do they turn to their smartphones in language production?*
3. *What are they relying on their smartphones for?*
4. *How do they use their smartphones to produce answers in their L2, English?*
5. *How often do they use their smartphones?*

I will continue by first explaining the context of the research. Then I will discuss the qualitative research methods that I decided to use, before focusing on particular issues that arose in their implementation.

The Research Context

As a part-time teacher at the university where I decided to conduct my research and with commitments elsewhere on other days of the week, I only had access to students for research on the day I taught there. The second-year students in the English communication class that I chose for this research were either business management or sports management majors. According to the university, 75% of students of their students had scored less than 380 on TOEIC, so they had low proficiency in English. The classes were compulsory, and my personal observation of students suggested a lack of motivation, especially among the sports management students. I assumed motivation would be greater if they were able to express their own ideas rather than only following vocabulary and grammar provided in the textbook. In order to find the words that best represented what they were trying to say, I encouraged the use of smart phones as dictionaries and for translation. It appeared that not only were students using simple word-to-word translation, but they were also translating complete sentences. The project design was threefold: choosing participating students, designing the task for participants to complete, and deciding which methods and techniques would allow me to collect the necessary data to answer my research questions.

Getting my students to become participants raised difficult issues, including the use and presentation of participants' personal information. Different ethical considerations are noted by Dörnyei (2007): shared information (Who has access to what?); collection methods (removing participants from their normal activities); anonymity (participants should not be named); handling the data (video data can remain after the project ends); and ownership of data (Who owns the data and controls the editing and releasing?) Barbour (2014) suggests that ethical issues should not only be a concern at the beginning of the project, but considered throughout the project. Another way of describing the need for ethics in research is a participant's exposure to harm. Traianou (2014) notes that harm can include anything from the

obvious physical harm, to harm by damaging reputations or projects participants are involved in. Some countries have laws in place to protect participants in research projects and to avoid exploitation, and require informed consent by participants, but Dörnyei (2007) notes that even when consent is not required by law, ethical principles cannot be ignored. Participants need to be aware that they are going to be involved in a study and should provide consent for this, as well as allowing the information produced to be used in a research project intended for a dissertation or for publication.

Selection of Participants

One of the most difficult decisions for my project was deciding how many students were necessary in order for me to obtain enough data, and how I would get them to participate in the project. This was important as it would impact the logistics, analysis, trends, and time constraints of the project. The more participants in the project, the more difficult it would be to find a time for all participants to complete the task. Additionally, the larger the number of participants, the larger the amount of data produced, and hence the more analysis required. Alternatively, would keeping numbers small give me enough information to discover trends in the data? Different to quantitative research where a large sample is needed to yield meaningful statistics, qualitative research needs to convey the individual experience by description and clarification (Dörnyei, 2007). There is no concrete answer as to how many participants are needed for a project. Dörnyei (2007) suggests six to 10 for an interview study, and if using computer–aided analysis, up to 30.

As this was my first time doing research involving participants, I made a blanket request to all of my students asking for five volunteers. There was no response. Unsure how to move forward, I contacted my supervisor; his suggestion was to approach possible willing students personally. Additionally, knowing that I could not increase grades for the helpers as ethical standards needed to be maintained, I was at a loss for how to reward those who gave up their time. Another suggestion from my supervisor was to give the participants a small treat at the end of the project. The selection process I used in choosing the five students was *purposive sampling,* which Padgett (2017) describes as "a deliberate process of selecting respondents based on their ability to provide the needed information" (p. 67). Each participant was chosen according to three criteria: firstly, their English score in my class in the first semester, secondly, the students' likelihood of saying yes, and thirdly, based on the possibility of their co-operation in doing what was required throughout all the tasks. I also ensured that there was a three–to–two ratio of males to females as a representation of the gender balance in the class, and to see if any trends could be observed for different genders. I decided to select students with different proficiency levels in order to find out if there were differences in smart phone use between ability levels, a decision based on my hypothesis that a higher level of English would positively correlate to less reliance on using the smart phone.

When I approached my chosen students, their two main concerns were: *What are you doing the project for?* and, *How much of my time do I need to give up?* I answered these questions clearly, and in a straightforward manner.

To ease any ethical issues that could arise I contacted the university department for any internal guidelines. A colleague suggested that if no names were used then there were no ethical considerations, but my instincts told me to go to the administration. As they had never had a similar request, samples of consent forms and an explanation of the students' roles in the project needed to be submitted. The students signed a consent form, ensuring they were aware the data would only be used for this research by me, and that at no time would their names be included in the paper. I also provided students with updates on the analysis of the

data to acknowledge their contribution. The five students whom I approached directly agreed to take part in the research project and gave up a half to full lunch period for completing the task and subsequent data collection.

The Task

The task designed for the students comprised three writing exercises. The first part was simple translation of Japanese words to English; the second part was an English question-and-answer style exercise; and the third part required students to write the plot of a movie. It was hoped that these three exercises would present tasks of increasing difficulty and would mean that students would use their phones in different ways to aid their completion of the exercises. As the level of some students was quite low, the task activities were explained in written Japanese.

Qualitative Methods Considered and Issues to Consider

Here I cover the qualitative methods I considered for studying the research questions: pre-task questionnaire, think aloud protocol, stimulated recall and interviews. I also consider, related issues that arose: triangulation and videoing. Most of these came from suggestions made by my supervisor, which I then further researched on my own. The importance of data validity is discussed by Dörnyei (2007), to ensure that criticism of the data collection is kept at a minimum, as no single method allows for this, I used a combination of different data collection methods. My research focused on the *why* of participants' behavior or reasoning, so I decided to use introspective methods, in the hope of obtaining data that would provide responses from unobservable thought processes. My first plan was to go with open-ended questions in a paper-based questionnaire.

Pre-task Questionnaire

In order to create participant background data, I administered a pre-task questionnaire in Japanese, students were able to answer in Japanese ensuring they could answer in detail and without hesitation, rather than translating. The questions focused on current English ability and participants' attitude to English learning and use. In order to capture the thoughts of why participants chose when and how to use their phone, I also used a post-task questionnaire asking the participants when they used their smartphones and for what purpose. The participants were required to answer the questions as soon as they had completed the task. I did not request the post-task questionnaire be completed as each individual activity was finished due to a concern for participants overthinking and losing their train of thought. The post-task questionnaire provided me with a first look at the participants' thought processes. Most of the answers were short and did not contain many details. For example, this was one answer provided by one participant when asked about their smartphone use: "I am not good at combining sentences." Some of the answers were difficult to understand even when written in the participants' first language such as "to make (write) it good." (うまく作るため). The answers from the post-task questionnaire were the basis for the semi-structured interviews, which I come to later in this paper.

Think Aloud Protocol

After consultation with my supervisor about the questionnaire, I looked further into his suggestion to use the *Think Aloud Protocol* (TAP) in order to generate data on the participants'

thought processes. This method encourages participants to verbalize their thoughts as they complete a task, providing a closer look into what they are thinking. Various suggestions have been made about how to successfully employ this method (e.g., Dörnyei, 2007; Ellis & Barkhuizen, 2005; Fonteyn, Kuipers, & Grobe, 1993) including the following practical points:

· participants need to be trained as vocalizing thoughts is not a natural practice;
· small sample sizes are needed due to the time required;
· individual sessions in quiet settings are necessary to reduce distractions;
· it is necessary to remind participants to verbalize thinking if pausing occurs to ensure all thoughts are vocalized;
· the data need to be transcribed at the completion of the session.

Additionally, the following points should be taken into consideration:
· the method needs to have no impact on the performance of the event (Dörnyei, 2007; Ellis & Barkhuizen, 2005).
· the method is more suited for writing tasks than oral tasks (Ellis & Barkhuizen, 2005)
· participants cannot pick and choose the information they verbalize, but must record all thoughts with accuracy for this method to be effective (Borg, 2006).

As this method seemed to me to be the most instant technique for recording thinking processes, I looked further into its application. I read papers that explained the method (e.g., Boren & Ramey, 2000; Charters, 2003), as well as papers that had used the method for data collection (e.g., Raimes, 1985; Wang & Wen, 2002). Through further research on Think Aloud Protocol (e.g., Dörnyei, 2007; Fonteyn, et al., 1993), I understood that training participants in how to speak their thoughts was integral to obtaining the necessary data. Before using this method with the participants, I tested its effectiveness in the early stages of my dissertation with a classroom conversation task. I instructed all of my students on how to vocalize their thoughts during conversations with each other. They were given an example of how it was to be carried out and prompted frequently throughout the task to express their thoughts. Although the students understood what was being asked of them, however, most of them failed to vocalize their thoughts. The students appeared more concerned with the task itself and having a conversation. As the students' English level is low, it was very difficult to apply out this method in English, and most students spoke in Japanese. Where I observed students using Japanese, it caused confusion with the task. I had not thought about the impact language could have on the process.

While debating whether to use this method in my dissertation, I noted that any vocalized Japanese would have to be translated. Students speaking in broken English did not verbalize all their thinking processes, and the quality of their English would make it difficult to understand what their thought processes were. I felt the effectiveness of the task would be affected if students were distracted by talking in Japanese. Most importantly, I could not be sure all thought processes were being verbalized, and I did not want to miss any important information. My observation was that when students focused on doing the Think Aloud as they were instructed, this appeared to impede the effectiveness of the lesson task. Another concern was that the natural flow of completing a task could be impeded by overthinking. Janssen, van Waes, and van den Bergh's (1996) study into the effects of the TAP in an L1 writing activity concluded that the reactivity of using the TAP varied according to the task, and that empirical checks should be used before applying the method. Even though I piloted this method with an oral task rather than a written task that I intended to use in the research project, my insecu-

rities and the training required for students outside of class time steered me away from using this method.

Stimulated Recall

As the TAP did not seem to be feasible for my research project, I needed to find another way to gain the necessary data that could be corroborated with the after-task questionnaire and expanded. My supervisor suggested I use *stimulated recall*. In this method, participants describe their thoughts retrospectively on an already completed task. The method can be used at intervals during a task through prompts such as *What were you thinking?* or using video playback of participants performing a task (Ellis & Barkhuizen, 2005). Dörnyei's (2007) recommendations to improve the output of stimulated recall include:

1. keeping the time between the task and recall within two days;
2. enabling participants to listen to and watch the task, as this provides more stimulation than written notes;
3. asking for thoughts, not explanations;
4. encouraging the participants to volunteer information rather than saying it for them, and
5. using participants' native language.

The first of these recommendations stems from the fact that our recall moves from short-term memory to long-term memory as time passes, and accuracy is compromised (Borg, 2006). Additionally, if the environment in which the recall is conducted is different to that of the "real-time" task, this may affect the information provided. However, stimulated recall provides an opportunity to express thoughts that would be difficult to obtain at the time of the task.

For successful implementation of stimulated recall, my participants needed to be video-ed while they worked on the written task. Videoing has the advantage of providing not only audio, but also visual data such as facial expressions (Ellis & Barkhuizen, 2005). However, it may be obtrusive and off-putting for participants, and may be ineffective if not set up correctly. Additionally, the transcribing and interpretation of non-verbal data can be difficult (Dörnyei, 2007). Three cameras were used in an attempt to view participants from different angles. As there was no need to collect spoken utterances, I did not have any concerns about microphone clarity and placement for recording what the participants said. The cameras were adjusted as the participants sat in their designated seats to capture their movements. While viewing the last participant working on their task through the camera lens, I found that the camera had been aimed at a higher sight line than anticipated. The student had hunched over their paper rather than sitting straight and valuable data had been lost. I adjusted the camera accordingly, but only clearly captured the last few minutes of them working on the task.

Once the task was completed, I needed to review footage of the five participants from the three cameras. I examined various aspects: discrepancies in the post-task questionnaire and the video evidence, instances of students in thought when answering questions, erasing and re-writing information, and using their smartphone. These observations were used to create questions for a post-task interview, as well as to design prompts for stimulated recall. Although participants had continual thought processes as they worked on the task, I was able to focus my prompts to the visual evidence of them in thought, leaning back on their chairs, looking around or at the ceiling in contemplation. Occurrences of participants in thought were recorded and time bookmarks were used for reference. Dörnyei (2007) notes that during stimulated recall interviewers should not ask leading questions, but rather should prompt the

interviewee to relive the moment by using questions such as *What were you thinking here?* As a week had passed since the task, most of the prompts for participants were *Do you remember?* Dörnyei (2007) also notes that interviews should be carried out in the participants' first language, and they were. The time needed to sort through the video was substantial. It took a full weekend to complete analysis of the five participants.

Once the analysis was completed, I was able to carry out stimulated recall with the participants. A week passed between administering the task and conducting the stimulated recall, well after the 48-hour recommended period. This was due to my teaching schedule at other institutions, and the logistics of getting to the university where the research took place. Additionally, students had other classes when I was available to talk, and lunch breaks were the only feasible solution. I knew that conducting all five participant sessions would not be possible in one lunch period, so I conducted some sessions during class time and the following week, 14 days after completing the task. Participants were shown video footage of the research task while having access to the task and follow-up questionnaire for recall. Whilst some answers provided usable data, participants often replied "I don't remember." One example was showing a participant video of them erasing content during the task whilst looking at their smartphone. It seemed odd that an eraser would be needed if copying from the phone. The prompt was *Do you remember?* Answer *Eraser?* Prompt *Why?* Answer *I made a mistake.* Prompt *Copying from the phone?* Answer *Probably I made a mistake when writing and I used the eraser.* Where participants could remember their reasonings and thought processes, valuable data could be collected. In addition to writing participants' answers, the stimulated recall was also video-recorded so as not to miss any information. The stimulated recall was administered in conjunction with a semi-structured interview so that more specific questions and issues could be raised.

Semi-structured Interview

An interview is an instrument for one party, an interviewer, to extract information from another party, the interviewee (Edley & Litosseliti, 2010). Interviews can be conducted in different ways, such as single or multiple sessions (Dörnyei, 2007), but usually consist of questions that require an oral response. In addition, interviews may be semi-structured or unstructured, individual or group based, and the delivery method may vary from face to face, to telephone, video phone, email, or chat rooms (Ellis & Barkhuizen, 2005). Interviews are not a foolproof method of finding out participants' thought processes for various reasons, including the time elapse between the task performed and the interview, as well as the inability to go back and ask questions that were missed in the first session.

In the semi-structured interviews, I asked participants both open and closed questions. The questions comprised fixed questions to all participants, as well as personally targeted questions. The specific questions for each participant were developed from queries about the answers in their post-task questionnaire, and discrepancies between the questionnaire and the video evidence. I was hoping the data gathered from the interviews would be the most detailed and useful in answering the research questions as I could see what I wanted to find out. An example of discrepancy between the video and questionnaire was a participant using their smartphone when they had not noted this on their post-task questionnaire. Going through each post-task questionnaire with each student allowed for clarification of their answers. One participant who failed to write reasons for use of their smartphone was able to provide the missing information. Another participant who had written "For how to summarize the synopsis" in answer to the question "Why do you use your smart phone to write the movie plot?" was able to explain their answer and this revealed that his intended meaning was different to

the one I had assumed. The participant had actually looked up an English version of the plot and was prepared to copy it word for word, rather than looking for a movie synopsis writing structure, as I had assumed. Interestingly, the students were quite willing to talk freely about their use of smartphones without any prompts.

Once the interviews were completed I transcribed each one and translated the relevant information from Japanese to English. This took a considerable amount of time as each participant conversed for at least 10 minutes, and some parts were difficult to hear. In order to include all utterances, multiple playbacks were necessary for certain sections, as was a native Japanese ear for the most difficult parts. The results were then written up and discussed in the dissertation. I analyzed the data by looking for common keywords. Students appeared to use their phones for similar reasons, and individual answers were expanded with the stimulated recall and interview. Findings from this analysis showed, first that regardless of level, when given the opportunity to use a smart phone for written English production, participants would readily do so. Second, the most frequently occurring use of smart phones was for translation purposes, including complete sentence translation for extended writing tasks rather than only unknown words. Additionally, the results revealed that some smartphone translations were simply incomprehensible, and some participants made errors when copying of information from their phones.

Triangulation

In order to ensure that my data were triangulated, I used a combination of questionnaire, stimulated recall, and semi-structured interview. Triangulation is a way of providing multiple perspectives on a phenomenon by using a variety of data sources (Denzin, 1978). In their paper on mixed method data collection, Moran-Ellis et al. (2006) suggest *integrating mixed methods*, where data generated by different methods are brought together, and provide an overarching view. Their approach incorporates the implementation of the triangulation methods, for example was the triangulation carried out concurrently or at separate times. They also note that using different methods "enriches the theoretical and epistemological approach" (Moran-Ellis et al., 2006, p. 20). Both views of triangulation suggest that by providing different vantage points of the same data we can create a more accurate picture of what is going on.

The use of video with participants for stimulated recall, exposed information that I did not know I was looking for. And using different methods and techniques enabled me to explore further into participant thought processes, and allowed for clarification of responses on the post-task questionnaire.

Post-Project Reflection

The following is a review of the various research tools used in my research project. As every project is different, the concerns raised here may not be applicable to other projects, but hopefully may help other novice researchers. In thinking over the different issues that occurred throughout the project, I became aware that one of my underlying concerns was time—and to have as much time as possible. Another reoccurring theme was translation of participants' thoughts and responses from Japanese to English. As a fairly competent non-native Japanese speaker, I was able to understand most of my participants' spoken Japanese; however, as I am not a professional translator, translation did not come easily, and extra care was needed for checking the accuracy of the translations I made. Overall, I believe my study did answer the research questions it posed. However, as the research was assessed as part of my MA, one of the assessor's comments was that some of the conclusions were

obvious. However, as there was little prior research in this area, I wanted to focus on covering the bases before identifying other areas of concern for further research.

Reflection on the Qualitative Research Project
Supervisor—Supervisee Relationship

The relationship with my advisor worked well as we both gave each other time. I didn't contact my supervisor every time I had a question, and he allowed me to work at my own pace in line with deadlines. The advice given worked better when researched further and piloted before the actual project. I also felt comfortable asking my supervisor for help when needed, the replies were always in a timely manner, and the information was not too detailed, but pointing me in the direction of where to find more explanation to promote self-learning. This scaffolding gave me more confidence to work on the project as I did not feel alone. If I were to do it over, I would have perhaps stayed in contact more regularly to gain feedback in smaller chunks. This would have allowed me to edit my research more frequently.

Selection of Participants

My initial disappointment at students not volunteering to be part of the project steered me to the use of *purposive sampling*. This allowed me to focus on students who seemed more willing to give information and participate earnestly. I still have concerns that purposive sampling did not provide a true representation of the class base, and it was researcher biased. Were students who were willing to participate in the study more motivated to work out a problem on their own before using their smartphone than students who did not seem motivated to participate? Had I chosen a balanced representation of those in the class? Although *purposive sampling* may not be a random selection of the population, it can address the issue of producing quality data by selecting participants who are communicative, available, and willing to participate (Etikan, Musa, & Alkassim, 2016).

Keeping the number of participants to five allowed me to find trends in the data produced, while keeping the amount of raw data at a workable amount. Each stage of the research project produced raw data that needed to be translated and transcribed before analysing. As qualitative research is not confined to giving participants a test and running responses through a statistical software program, it can be very time consuming (Dörnyei, 2007).

The decision on how to reward the participants for their time was also difficult. Whilst little of the literature I had read gave suggestions on how to thank participants, the advice from my supervisor to gift some sweets seemed appropriate and was well received. Additionally, keeping the participants informed of how and where the data were being used, as well as the conclusions reached in the project, maintained participant motivation and appeared to give the participants a sense of accomplishment.

The consent of not only the participants, but also the institution where the research was carried out, was needed to maintain ethical standards. It surprised me that although other professors at the university carried out research using students as participants, the institution did not have any official procedures in place. While I was informed that student consent would be sufficient, my pressing for institutional consent ensured that procedures are now in place for others to follow.

I would be very likely to choose my participants for another research project using the same procedure without changing anything. Choosing students who appear to be willing and capable ensured that I was able to gather enough data for analysis.

Post-Task Questionnaire

It needs to be noted that the post-task questionnaire was the only method applied within a short time frame of the task being administered. It allowed for short-term memory recall by requesting information through open and closed questions. This method was the basis for the follow-up interviews and allowed me to develop questions to probe deeper in the interviews. The responses given in the questionnaire were simple and appeared mainly in note form.

When thinking about the effectiveness of the questionnaire, I have the sense now that it would have been beneficial for participants to be aware of what questions were to be asked about how they completed the task. Alternatively, participants could have been working with the questionnaire as they completed each section in the task to reduce the time between task and question.

I would certainly use this method again as an instrument for obtaining data on thought processes. As it appeared that the questions were not stimulating enough for participants to give detailed answers, I would however need to change the way the questions were formulated. A simple *Why?* did not encourage long responses. As it is difficult to know exactly what each participant is going to write, it might be difficult to preempt all questions, although a pilot may identify some issues.

Think Aloud Protocol

Whilst the literature noted the think aloud protocol (TAP) was more suited for written tasks, I piloted this with a conversation task to see if students were able to voice their thoughts. This revealed that the task was confusing for students, which led me to decide not to use it for this research project. Had I piloted it with a written task, the results might have been different and more suited to the research project.

It was clear from the piloting that student training needs to be undertaken for it to become a natural process. This needs to be carried out over a period of lessons and a variety of situations for greater effectiveness, rather than a one-off as used in this project.

The pilot also made me aware that transcribing and interpreting the data would take time. Another more difficult aspect involved transcribing the students' utterances into usable data if the students were not speaking coherently. This would require interviewing students about their thought processes within a short time frame to confirm what they had been thinking at the time they were doing the activity.

Should I attempt a written research project with students again, I would like to retry this method as a way of glimpsing into students' thoughts. I would also do further research into training students to use this method, such as the training methods described by Fonteyn, et al., (1993), and van Someren, et al., (1994). In addition, I would ensure that the all the recording devices were well tested and positioned before starting any tasks. I do believe that this method alone would not suffice in gaining necessary data and would need to be supplemented with another method such as a questionnaire, stimulated recall, or interview.

Stimulated Recall

As all the participants completed the task at the same time, it was impossible to do the recall session with all participants at the same time. Rather than administering the research task to all participants simultaneously, having each participant complete the task individually followed by stimulated recall may have proved more fruitful. My research into this method before implementation had not prepared me for the repetitive *I don't remember* response to questions.

Another complication of the session was its combination with the semi-structured interview. With students needing to attend other classes, the time frame available to work through both instruments resulted in my rushing of the video recall. Rather than giving participants the time to recall what they were doing, I felt after a short period of silence a necessity to move on.

I would be very interested in using this method again, providing it was carried out much closer to the completion of the task. As prompts for participants are not specific, there is little need to prepare, and if carried out one student at a time, I could use my researcher observations in the moment as prompts without reviewing the video. This would let me keep the task and recall within a tighter time frame.

Interviews

The interviews allowed for depth to be added to the post-task questionnaire. The questions took time to prepare, as various methods had been used to collect data up until this point. The interviews were videoed for ease of transcription later, which let me focus on the conversation rather than note taking. The interviews were also combined with stimulated recall. As the participants found it difficult to express their ideas in English, it was necessary to conduct the interviews in Japanese, my second language. There were points in all five interviews where misunderstandings occurred either due to language barriers or misunderstanding of concepts. These were rectified on the spot by either myself or the participant. Some interviews were longer than others, as some participants presented more information than others. The longer interviews provided more depth and generated new questions to be answered. I was disappointed that I was unable to conduct the interviews much closer to the task being administered. There was a definite lack of information from participants when I asked them for their explanations of certain answers or activities in the task. Obviously, time was required to carry out the individual interviews, but the consequent transcribing and translation of the interviews also took a much longer amount of time than expected. The interviews enabled me to confirm data that had been presented in other forms already, and clarify ambiguities. Notably, my interpretation of a response to a question by a participant had a significantly different meaning than intended, and allowed me to amend my analysis of the data. I believe this was the most successful method of finding out why participants had done the task in a particular way. I would have preferred more time to conduct the interviews without encroaching on participants' study time or personal lives, as I believe this would have generated deeper information.

Due to the effectiveness of this method in this research project I would definitely use interviews in future research. As with the stimulated recall and questionnaire I would aim to administer the interview much sooner than in this project and at a more convenient time for the participants.

Videoing

As noted previously, videoing allowed me to see things that were not visible at the time of the research task implementation. I could see when the participants moved or appeared to be thinking and this prompted specific questions for the interview and stimulated recall. The footage of five participants from three cameras took longer than expected to review. As the time constraints of analyzing video and carrying out the interviews with five participants made it impossible to complete the stimulated recall within 48 hours, I would likely attempt this again by carrying out the task individually, one on one with a single participant, so that I could review the video and interview each participant in a timely manner. This would allow

participants to recall information more easily, but increase the overall time needed for participants to complete the task one by one. As a trade-off for the extra time needed, I would expect that the depth of the information recalled would create more useful data. I would also complete a trial video recording of tasks in action to ensure that the video angle captured the participants' behaviour fully and clearly.

Triangulation

As noted above, by not relying on one method, I was able to strengthen the data used in the project. Written information was checked in the interviews for any misunderstandings and video provided evidence of physical movement which could be checked with the written responses. As the research task was a written task completed in silence, the use of video to corroborate participants' written responses gave an added dimension to the research. I was unaware how useful this data would be until I watched the video with the task responses and after I had the task questionnaire in hand. I was surprised at the number of discrepancies between participants' post-task responses and what had taken place. The most obvious was student's failure to note smartphone use in their post-task questionnaire when the video clearly showed they had. When doing further research, I intend to carry out video tasks, as well as follow-up interviews and questionnaires to create a complete picture.

Overall Observations

Regardless of the any techniques mentioned earlier, the most critical reflection I take with me is the notion of time. Bell & Waters (2014) observe that "the extent of your data collection will be influenced by the amount of time you have" (p. 120) and I would fully concur with this. The importance of time has been a common thread through the evaluation of each method and technique. The importance of a short time frame for short-term memory recall was made evident in the interviews. Additionally, the unforeseen time needed for translating and transcribing made the project longer than expected. This too is the most important advice for someone new to qualitative research: be aware that everything takes longer than you expect.

Being able to combine different qualitative methods allowed for data to be crosschecked and triangulated. If only one method had been chosen, the data would have only been taken at face value. Finding the ideal way for the methods to complement each other is also important, as each method allows for slightly different information. The video evidence revealed inconsistencies between what the participants actually did and what they remembered when answering the questionnaire.

Researching the methods and their applications can not only save a lot of time, but it can also help you to find more effective ways of obtaining your required data. Take suggestions from friends and colleagues, and research them to see what fits your project best. Piloting the research techniques and methods that you intend to use will allow you to see inconsistencies and will permit you to make corrections, developing the way you research as you learn from experience.

Conclusion

In this paper I have presented a description of and reflection on the implementation of qualitative research methods. As the research project was carried out as part of a Masters degree under supervision, it may differ slightly to that of "stand alone" research for a journal publication. For future research I would discuss the process with colleagues in order to gain advice and insights about my project, mirroring the work of a supervisor. The opportunity to revisit

this research project and work through the techniques has been beneficial for me and allowed me to reflect on the decisions I made at the time as a novice researcher. The process of working with an editing group for writing this paper for a journal has also been a new experience that has enabled me to grow as a researcher. The time and effort required for making multiple reviews over several drafts was something I was not aware of. I have also been able to assess the effectiveness of each method and technique, enabling me to see not only flaws in the method itself, but also my lack of knowledge in each method, as well as the challenges. Reflecting on each tool has allowed me to clarify how I can use these methods more effectively in future research projects, including how students use smartphones when overseas and the impact that would have on language learning. I learnt that the restraints with time and logistics were big factors in the selection of my methods, as well as their effectiveness. Not being able to gain access to students in a timely manner after carrying out the research task meant that vital participant thought processes were lost. In future research I intend to be more conscious of this issue. Although confident that I had done enough research to use the methods effectively, errors in implementation and loss of critical data show a need to do more background reading of recommendations and work by other researchers. This will allow me to collect data that would yield more substantial findings.

Reading the theory of how a particular method works and should be implemented does not necessarily correspond to implementation and effectiveness in any research project. The suggested way of execution is not always the way it turns out. As each research project is different, what works well for one researcher may not work so well for another. Reading available literature on the methods, while extremely valuable, will not necessarily lead to success in your chosen research.

This research project allowed me to see a little further into the thoughts of learners, that simple word translations were not the only way to use smart phones, and that students' confidence in technology outweighed confidence in themselves. I saw the participants working hard in their free time to help my project and completed the tasks to the best of their abilities although not being given any credit. Additionally, I learnt that, even with an advanced tool such as a translation app, learners still need to be taught how to use it effectively, and how to recognize mistakes in the English that they produce. It appeared that the learners were content to rely on such technology regardless of its effectiveness.

By the end of my research project and confirmed in my writing of this paper, I came to understand the importance of time in a qualitative research project, and that there is never enough.

References

Barbour, R. (2014). *Introducing qualitative research: A student's guide* (2nd ed.). London: Sage.

Bell, J., & Waters, S. (2014). *Doing your research project: A guide for first-time researchers.* Maidenhead: Mc Graw-Hill Education.

Boren, M. T. & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. *IEEE Transactions on Professional Communication, 43*(3), 261–278.

Borg, S. (2006). *Teacher cognition and language education.* London: Bloomsbury.

Charters, E. (2003). The use of think-aloud methods in qualitative research An introduction to think-aloud methods. *Brock Education Journal, 12*(2), 68–82.

Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods.* Englewood Cliffs, NJ: Prentice Hall.

Dörnyei, Z. (2007). *Research methods in applied linguistics.* Oxford: Oxford University Press.

Edley, N., & Litosseliti, L. (2010). Contemplating interviews and focus groups. In L. Litosseliti (Ed.), *Research methods in linguistics* (pp. 155-179). London: Continuum.

Ellis, R., & Barkhuizen, G. (2005). *Analyzing learner language.* Oxford: Oxford University Press.

Etikan, I., Musa, A. S., & Alkassim, S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics, 5*(1), 1-4.

Fonteyn, M. F., Kuipers, B., & Grobe, S. J. (1993). Description of think aloud method and protocol analysis. *Qualitative Health Research*, *3*(4), 430-441.

Janssen, D., van Waes, L., & van den Bergh, H. (1996). Effects of thinking aloud on writing processes. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 233-255). Mahwah, NJ: Lawrence Erlbaum.

Moran-Ellis, J., Alexander, V. D., Cronin, A., Dickson, M., Fielding, J., Sleney, J., & Thomas, H. (2006). Triangulation and integration: Processes, claims, and implications. *Qualitative Research, 6*(1), 45-59.

Padgett, D. K. (2017). *Qualitative methods in social work research.* Los Angeles: Sage.

Raimes, A. (1985). What unskilled ESL students do as they write: A classroom study of composing. *TESOL Quarterly, 19*(2), 228-258.

Traianou, A. (2014). The centrality of ethics in qualitative research. In P. Leavy (Ed.), *The Oxford handbook of qualitative research* (pp. 62-80)*.* Oxford: Oxford University Press.

van Someren, M.W., Barnard, Y.F., & Sandberg, J.A.C. (1994). *The think aloud method: A practical approach to modelling cognitive processes.* London: Academic Press.

Wang, W., & Wen, Q. (2002). L1 use in the L2 composing process: An exploratory study of 16 Chinese EFL writers. *Journal of Second Language Writing, 11,* 225-246.

Willig, C. (2013). *Introducing qualitative research in psychology* (3rd ed.). New York: Open University Press.

Review Process

This paper was peer-reviewed by Daniel Hooper. It was also blind peer-reviewed by the members of the Learner Development Journal Peer Review Network. (Contributors have the option of open or blind review.)

Author Bio

Clare Kaneko is a visiting instructor at Keiwa College, Niigata. She earned her Master of Arts TEFL/TESL in 2017 from The University of Birmingham, England. Her research interests are CALL/ MALL, and classroom materials. She is currently expanding her research knowledge with an extensive reading project using quantitative methods, hoping to compliment her qualitative knowledge. Email: kanekoclare@gmail.com