Transcribing With ATLAS.ti

Rohaiza Zakaria, Aede Hatib Musta'amal, Nor Fadila Mohd Amin, Halmi Mashindra Saleh

Abstract

Most qualitative studies engage interviews and/or observations in their data collection phase, which comprise of audio and/or video data. These data are usually transcribed into a written form for further analysis. Since transcribing takes considerable amount of time, an efficient tool is required to prepare for the transcription. In fact, there are many software available in the market to facilitate transcription including Inqscribe, Express Scribe and F4 Transkript. These transcription programs, however, were developed solely for transcribing purpose. ATLAS.ti, in addition to its superior tool set for the analysis of qualitative data, also can be used for transcription. In this paper, the authors emphasize the importance of the transcription process in qualitative studies, and share their experiences in using ATLAS.ti for transcribing audio and video data for a design study project.

Keywords

Introduction

Our engagement with the issue of transcription was initiated by our own research project, in which we were identifying sketching behaviors among student-designers. We met participants twice a week to record the process during their Design Project classes. One observation session normally took about two to three hours, followed by a 30-60 minutes interview session. This resulted in a huge amount of video and audio data for almost every session. As the study was inductive in nature, we transcribed those files immediately, right after the observation or interview sessions; or at least, transcription was completed before the next meeting with the participants. This was done in order to ensure that we could perceive and understand the data clearly. If there were doubts that needed to be clarified, we could ask the participant during the next field visit.

It is only normal for qualitative researchers to engage with a great deal of audio and video data, and in order to analyze them, the researchers can simply listen to the audio recordings or view the video data. However, this process only enables a superficial level of analysis. In order to achieve a thorough, ongoing and deep analysis, the interviews have to be transcribed into a text documents (Wilson, 2002). Transcription is considered a first step in managing and analyzing recorded data; it is a representation of audible and visual data into a written form (Bailey, 2008; Evers, 2011). In other words, it is a written record of the interview or observation (Arksey & Knight, 1999).

Why Is It Important To Transcribe Data?

Transcribing is a difficult and time-consuming task that most qualitative researchers dislike (Alcock & Iphofen, 2007; Matheson, 2007; Wilson, 2002). A majority of researchers try to avoid transcription altogether, or hire transcriptionists to do the job for them. This, however, could lead to ethical issues and endanger the credibility of the findings (Easton, McComish, & Greenberg, 2000; Tilley, 2003; Tilley & Powick, 2002).

Transcription is a very important aspect in qualitative inquiry which should not be skipped (Oliver, Serovich, & Mason, 2005). Novice researchers are often unaware of the fact that transcribing is itself an analytical process that can influence the research study in significant ways. The transcribing process is valuable and can yield important insights (Lapadat, 2000). It should not be seen as merely a mundane task in the research process because it can be directly related to findings of a particular study (Skukauskaite, 2012) and it can be seen as a form of analysis too (Hammersley, 2010). Moreover it involves close observation of data, which allows the researcher to grow closer to and achieve greater familiarity with the data (Arksey & Knight, 1999; Bailey, 2008; Lapadat & Lindsay, 1999; Tilley, 2003). It also provides the opportunities to listen to the participant's words, pauses, silences and non-verbal expressions as well. It is vital in qualitative research to have an in-depth understanding of the data. As the analysis and understandings are derived through listening and re-listening to the audio (or viewing and re-viewing the video), transcription facilitates interpretive thinking, which is needed to make sense of the data too (Bailey, 2008; Lapadat & Lindsay, 1999) .

Furthermore, transcribing also gives researcher the opportunity to grasp and spot evidence better and more clearly. Voice intonations or even body language can also be interpreted in many ways (Mishler, 2003). The addition of pauses, laughter, and body conduct to a transcript invites a different interpretation of a situation. Representing some non-verbal features of the interaction in the transcript can change the interpretation of the interaction. Thus, it is very important for researchers to go through the process of transcription personally to ensure that every detail of data can be well acknowledged. This familiarity with data and attention to every detail that occurred during data collection can facilitate realizations or ideas which may become useful during data analysis (Bailey, 2008; Lapadat, 2000).

While transcribing, researchers will automatically reflect on the on-going research strategy, and at the same time it can help them to improve their data collection method. If there are any questions that have been missed out during a session, or new ideas arise, these questions and ideas can be asked and implemented during the subsequent sessions. Miles, Huberman, and Saldana (2014) also emphasize the concurrency of data analysis and data collection. Transcribing helps the researcher to cycle back and forth between thinking about the existing data and generating strategies for collecting new and better data. It provides a unique opportunity for researchers to critique their own work and to potentially improve upon their interviewing techniques as well (Johnson, 2011).

In our case, this dialectic process of data collection, transcription and initial analysis has improved the quality of the collected data as we are able to reflect upon the data collection process during transcription. When playing back the audio and video recordings while transcribing, we automatically reviewed what we had done during the previous data collection session. As a result, we improved the way we conducted interviews (e.g., not interrupting the interviewee until he or she finishes talking). We also learned how to best handle the video camera while recording participants' sketching behavior. Over time we

figured out from which angle to record in order to facilitate transcription and interpretation of the participants' behavior.

Finally, a very concrete reason for transcribing is the current journal format, which is mostly printed text. At present, journal and thesis publication formats are mainly in pdf. This format does not permit multi-media files to be embedded. Furthermore, based on ethical reasons it is important to ensure anonymity of the participants when disseminating reports or even during presentations (Arksey & Knight, 1999; Creswell, 2013). Therefore, the best way to present the findings is to transform the audio and video data into text for reporting purposes.

With all the reasons that have been mentioned, it can be concluded that transcribing is indispensable in qualitative research.

However, transcribing is a time-consuming activity with at least three hours of transcribing per hour of participants' talk, and up to ten hours of transcribing per hour of video play with a fine level of details, including visual details (Arksey & Knight, 1999; Bailey, 2008). It is not an easy task. Many researchers struggle with the task of transcribing because they see it as a tiresome, stressful, lengthy, and challenging process that requires specialized skills, patience, and physical ability (Arksey & Knight, 1999; Lapadat & Lindsay, 1999; Tilley & Powick, 2002); some types of study even can cause an emotional impact to the researchers (Kiyimba & O'Reilly, 2015).

It is an undeniable fact that transcription is a challenging process (Tilley, 2003), especially when the researcher has to deal with hundreds of audio and/or video recordings. Normally, we had roughly about 2-3 hours of video recording, and one hour of audio recording to be transcribed from every observational session conducted that needed to be transcribed within two days before meeting the participants for the next session. This was not an easy job. For novice researchers like us, we initially transcribed using the human transcription method, where we kept switching back and forth between the media player and the text editor to transcribe our data, until recently we were exposed to several dedicated transcription tools. These tools have simplified the transcription chores and sped up the process.

Which Tools Can Be Used?

There are many transcription programs available that can aid the researchers in transcribing, such as F4, Transcriber, Express Scribe, and Transana. These transcription programs basically facilitate the manual transcription process. They have features that allow both text editor and audio/video player to run on the same screen, and the digital recordings can be managed without hassle. However, they are merely meant for transcribing purposes.

Meanwhile, CAQDAS such as ATLAS.ti, MAXQDA or NVivo also provide transcription tools for researchers while also facilitating researchers in managing their qualitative data. The transcription tools in CAQDAS make the qualitative data analysis and the management of data easier. This is because all the

data are kept in a single software, where researchers can transcribe and later carry out the analysis in a single location (Evers, 2011). Hence, researchers do not have to buy both types of software licenses. This economic reason is one of the advantages that CAQDAS offers.

Among the many CAQDAS that are available, we selected ATLAS.ti for preparing transcripts and for analyzing the data. The program has many great features that served the purpose of our research project well—besides managing and facilitating the process of analyzing qualitative data, ATLAS.ti also offers a built-in transcription tool. While this tool has similar features to other transcription programs, it offer a number of additional advantages which are further elaborated in the next section.

Transcribing With ATLAS.ti

The Process

Both audio and video files undergo the same process of transcription. There are three stages in transcribing audio and video data with ATLAS.ti., namely: adding files, associating multimedia and text file and transcribing. Every stage is followed by several steps. The flow chart of the process is as shown in Figure 2.
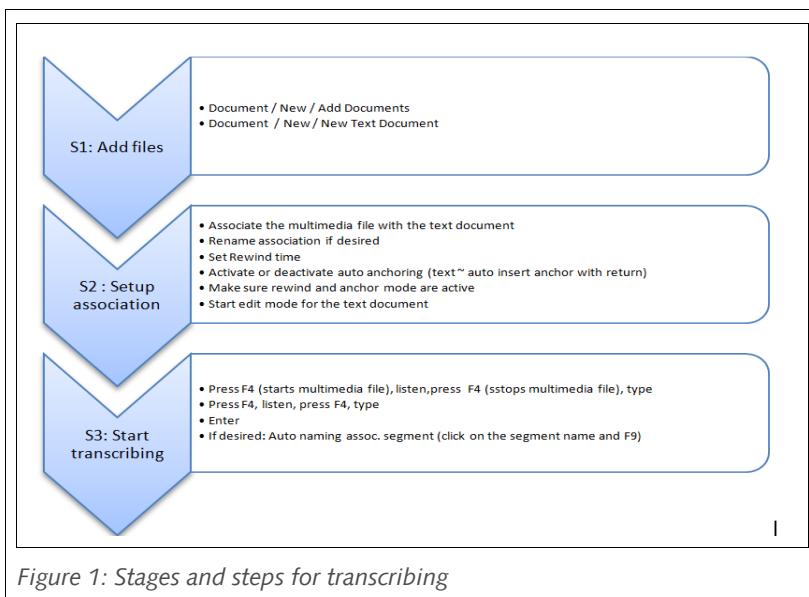

Figure 1: Stages and steps for transcribing

Preparing Documents For Transcription

In order to start transcribing, researchers have to add the audio or video files to an ATLAS.ti project. Next, a new text file needs to be created and associated with the multimedia file. The text file and the multimedia file can be displayed side-by-side, so it is easy to manage both documents – and researcher can simultaneously view and work on the transcription process (see Figure 2).

When associating the multimedia file with the transcript file, the Association Editor opens with a default name. Researchers can rename the association for easier reference, e.g., `Hafiz intw transcript PO1`, `Roy vdi transcript 20130725`, and so on.
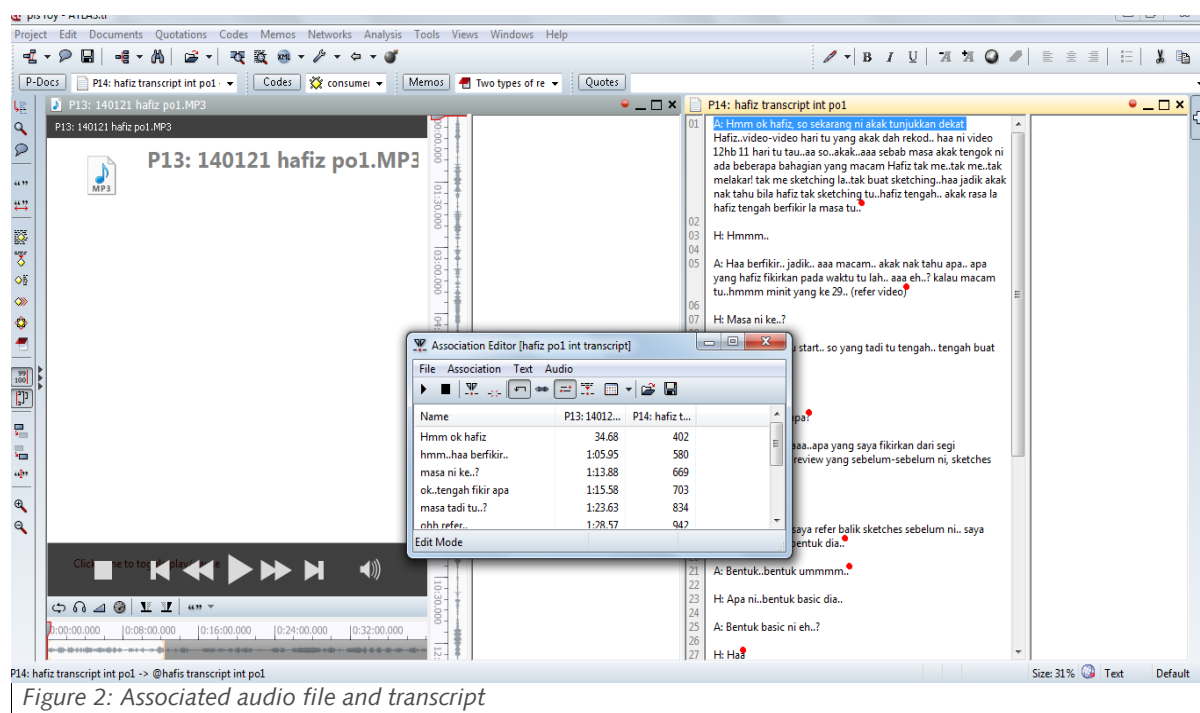
Figure 2: Associated audio file and transcript

One audio or video recording can be associated with multiple transcripts. This is useful if you want to produce different types of transcripts, e.g., a verbatim transcript and a more detailed Jeffersonian transcript for selected parts of a recording.

Another useful option is to set the rewind time in the Association Editor, where researchers can choose the interval time to rewind the audio from 1 to 5 seconds. This 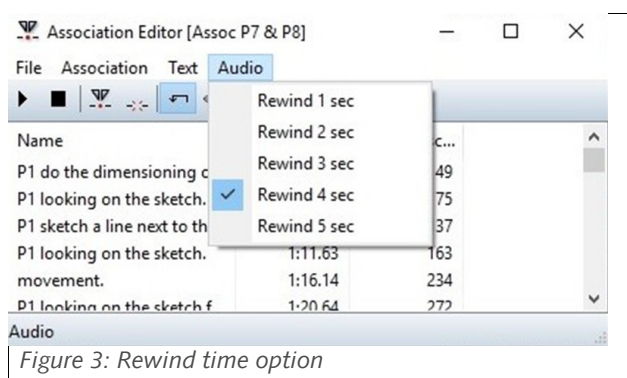function allows the researcher to listen or view the part that has already been transcribed based on the selected time interval (see Figure 1).

Figure 3: Rewind time option

Another option is to set time stamps, either manually or automatically. In ATLAS.ti, they are called *anchors*. If you opt for the automatic option, a time mark is set at the end of each paragraph, or in technical terms; every time you hit the ENTER key. This is a bit more convenient and saves one mouse click. The synchronization of transcription and multimedia file, however, becomes more accurate if you insert anchors more frequently. This is useful if you want to continue to work with the associated documents for further analysis purposes. For preparing textual transcript, it is sufficient to have time marks at the end of each paragraph.

Before starting to transcribe, double-check whether the REWIND MODE icon and the ANCHOR icon in the Association Editor are activated. In addition, the transcript must always be in EDIT mode. This is automati-

cally the case after creating a new text document. When, for example, taking a break and closing and restarting ATLAS.ti, all documents will be in read-only mode by default. Thus, you need to remember to click on the EDIT button in the main tool bar before you can continue to work on your transcript in a new session. Another important issue is to know that saving the project file is not sufficient—you first need to save your transcript and then your project file. Usually, saving the project is not possible without having saved the transcript, but computers tend to crash once in a while. Therefore, we recommend saving your transcript regularly, e.g., every 20 minutes by clicking on the EDIT button in the tool bar. Select SAVE ONLY. Next, save the project file (i. e. click on the disk button, or use the key combination CTRL+S, or select PROJECT / SAVE from the main menu.

The Transcription Process In ATLAS.ti

When everything has been set up, researchers can start transcribing the multimedia file. To start and stop the multimedia file, you can either press the F4 function key or use a food pedal. The foot pedal needs to be configured with the F4 key in order to ensure it works with the ATLAS.ti transcription tool. The ad-vantage of using a foot pedal is that the researcher has his or her hands free to type, without having to click on the F4 button. This definitely speeds up the process.

After pressing the F4 key, start listening to the recording, e.g., to half of a sentence or as much as you can remember. Press F4 again to type what you have just heard. If you prefer to manually set time marks, press the F8 key or the corresponding button in the Association Editor. Time marks are shown in the Association Editor as red dots in the transcript. If you want to name each anchor, this can be done by clicking F9. ATLAS.ti then inserts the last line of the segment as anchor name (see Figure 4). The re-searcher can later use these anchors to navigate through the document.
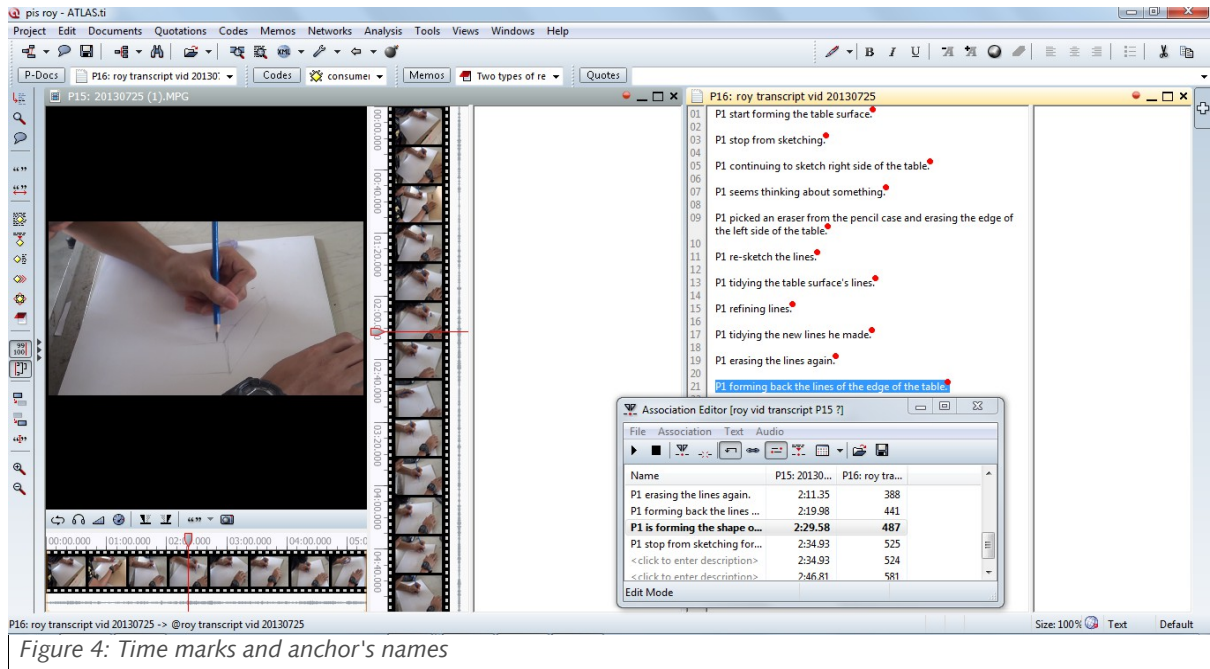
Figure 4: Time marks and anchor's names

Two further options are worth mentioning: The replay of the original recording from any position in your transcript, and the "Karaoke" mode. Let assume that you have already coded the transcript and want to listen to a particular data passage, this is what you need to do: Make sure the syncro mode is activated for the document (press the F3 key), select the coded segment press the key combination Ctrl+P (this option can also be accessed via the main menu). This loads the associated multimedia file (if not already loaded) and the part of the file that corresponds to the selected piece in the transcript is played back. The more time marks you have set, the better the synchronization will be, i. e. the precision of starting and end points of the replayed segment. In "Karaoke" mode, the multimedia file is played and at the same time the corresponding transcript is highlighted. For more detailed technical instruction, see the official program manual.

Based on our experience, the various function keys have greatly simplified the process of transcribing and eventually saved precious times that we could use to focus on other research matters.

What To Transcribe?

It is common practice nowadays that researchers use audio and video recorders to record their inter-views, observation sessions, and even their personal research diaries. The use of video recorders is gain-ing popularity among researchers recently due to technical advances and ease of use (Heath, Hindmarsh, & Luff, 2010). New technologies provide researchers with rich and detailed data about individuals' own interpretations, perceptions, experiences and practices (Wright & Russell, 2010).

Decisions related to transcribing are guided by the methodological assumptions a researcher makes in the context of a particular research project (Bailey, 2008; Skukauskaite, 2012). Transcripts can vary in level of detail. Researchers can choose to prepare a pragmatic, a standardized Jeffersonian, or a gisted transcript (Evers, 2011). A pragmatic transcription may be verbatim, but does not follow any fixed rules or guidelines. A Jeffersonain transcript captures in detail speaker turns, overlapping speech, pauses, tone of voice, pace, etc. It is the most time consuming transcript to prepare. A gisted transcript in comparison only contains summary phrases.

When working with video data, researchers may decide to transcribe the verbal and the non-verbal data like facial expression, body language, gesture, posture, laughter, background noise etc. Heath et al. (2010) suggest to also transcribe contextual information such as the material environment, participation and institutional practice. This is also known as description transcription (Demster & Woods, 2011; Paulus et al., 2014).

The possibility in ATLAS.ti to associate multiple transcripts with one media file nicely supports the parallel use of various transcription formats. When transcribing video data, one may want to prepare a verbatim transcript for what can be heard on the recording and a description transcription for the non-verbal interaction (Evers, 2011; Paulus et al., 2014). Researchers also may want different types of transcription for one audio file, e.g., a gisted transcription for the whole file and a Jeffersonian transcription for certain segments only

Tip: When transcribing audio or video data and a speaker provides long answers, it is recommended to divide speaker units into multiple paragraphs. Add at least one empty line between two speakers to indicate a speaker turn. This facilitated later reading of the data, makes automatic coding in ATLAS.ti easier and also ensures accuracy of synchronization between the transcript and the media file.

Another choice when working with videos could also be to transcribe only the verbal communication, and to *code* all non-verbal interactions. Given that each coded segment ("quotation") is an individual object in ATLAS.ti, the quotation name can be used to add short titles to each video quotation and the comment field to summarize what is going on in the video segment. Codes can be used for the next level of abstraction, i. e., for summarizing video quotations by theme, type, or category.

A possible option is to forego transcription altogether and to segment, comment and code audio and video data directly. On the surface, skipping the transcription process will save a lot of time. Evers (2011) also found that students in her classes felt much closer to the research participant as they could hear the intonation while analyzing the data instead of just reading a transcription. Similarly, Wright and Russell (2010) proposed that researchers are "sensorially" closer to the data when directly coding the multimedia files instead of working on the textual transcription of the same data.

The down-side is that direct coding tends to induce a premature analysis as this process goes too fast and there is insufficient time for data immersion for researchers. Evers (2011) reported that the elimination of

the visual stimulus of reading had affected the analytic thinking of her students. They did not have sufficient time to reflect on what they heard as the process of sub-vocalization is missing when not reading a transcript. When we read, we hear in our minds the words we read, and this brings us back to the interview situation, which in turn elicits our sensorial memory. Evers (2011) also found that researchers easily became sloppy in rephrasing the respondents' speech and forgetting whole parts of the dataset after the initial analytic round. Overall it took more effort to code audio or video data directly than to read a transcript, as it is much easier for the eye to move back and forth across a printed page than to stop and rewind an interview (Wright & Russell, 2010). Evers concluded that direct coding is suitable for broad coding only but not for fine coding. She recommends to use direct coding for contextual information, i. e. to prepare a description transcript, and not for verbatim interactions.

Our point of view is that if a project requires researchers to study behavior or interaction, you may consider transcribing the observational part as well because the analytical process that occurs during transcribing is important and impacts the interpretations to be made. It gives opportunity for researchers to immerse in the data and to avoid premature analysis. If the purpose of recording video data is only to provide additional information, you may consider the code the video directly.

In deciding whether to prepare a transcript or not, two further issues need to be considered: the dissemination process and ethical considerations. Most current journal formats do not allow to embed multimedia data into an article. Researchers still have to convert the media file segments into text in order to make it possible to be published. A possible work-around is to work with sequential images. And last but not least, ethical consideration might also speak against direct coding of media files. Preparing a textual transcript is often the only way to ensure complete anonymity and confidentiality (see also Creswell, 2013). [1]

Advantages Of Using ATLAS.ti For Preparing Transcripts

As we were going to use ATLAS.ti for analyzing our data, it was already a first advantage to be able to use it for the transcription process as well. This meant that we did not have to learn nor to buy another software. Furthermore, we could already begin with our analysis while transcribing due to the integrating memo and comment function. Whenever we had a hunch or emergence of insight while transcribing, we could immediately record it in a memo. As Johnson (2011) points out, such hunches and insights that

---

1 Editors' note: The time saved by not transcribing audio data may indeed be offset by the longer time it takes to set start and end position of audio quotations as compared to coding textual transcripts. However, the described danger of premature and sloppy analysis can be countered by a more intensive usage of the quotation level, especially when working with video data. The quotation level is highly under-utilized in ATLAS.ti and often overlooked. Instead of quickly adding a code, one could first spend some time to segment the data; that means creating quotations, name them, and write comments on the quotation level before one moves on to code the data. This is likely to be more insightful than spending a great deal of time transcribing non-verbal and contextual information. Interesting data segments can immediately be linked via hyperlinks, and one is already in the midst of the analysis rather than still being occupied with the transcription. For dissemination purposes, one cannot only draw on the original multimedia data but also on the text produced in form of quotation names, comments and hyperlinks.

emerge during the transcription process are very important to record as they may help in scaffolding the theory building.

Using ATLAS.ti from the very beginning in our research project, we recorded methodology reflections, research insights, wrote our research diaries and field notes using the memo function. At times, when transcribing our data, we noticed that our interview skills needed a bit of improvement, e.g., we interrupted participants while they were still talking, and some of us used to stutter a lot in the conversation, too. So we collected all of these observations in a memo and reviewed it before the next interview session, so we could do it better the next time. These kinds of reflection were part of our learning process during data collection, and they helped us to refine our methods and strategies in conducting the research.

Whenever we noticed and learned something when transcribing the data that was relevant to our project, we recorded those in a memo. Sometimes, we also captured situations that happened during an observation or interview session that was not recorded but may have a potential effect on participants' behaviors (e.g., a lecturers' instructions, an unpleasant environment, etc.) so that we could keep track of those situations. In sum, the ATLAS.ti memo function helped us to keep an overview of the data collection progress and allowed us to gain first insights while transcribing our data. This definitely had a strong influence on our continuing data interpretation and provided a framework for latter theory building. Memo and annotation functions are only available in CAQDAS, such as ATLAS.ti, which permits to record ongoing ideas and first insights already during the transcription process (Evers, 2011).

Another advantage is that ATLAS.ti allows to transcribe both audio and video data. Some transcription programs only support the transcription of audio data. Since we had both audio and video data, we gained experience in transcribing both data file formats and found it quite comfortable. The screen is split in two parts: On the right hand side the multimedia file is shown, and the margin area on the right hand side shows quotations, codes, indicates where comments have been written, shows hyperlinks and segments with attached memos. The timeline runs from top to bottom and makes it easy to track the data. Furthermore, the timeline can be zoomed, so that it is also possible to work comfortably with longer recordings.

We also found it quite useful that we could immediately attach codes when we noted some trends or pattern during the process of transcription. Hence, the transcription and coding process sometimes took place simultaneously.

The Future Of Transcription

With the current rapid advance of technology, there are increasing numbers of researchers that have started to explore and use Speech Recognition Software (SRS) in reducing the typing chores (Al-Aynati & Chorneyko, 2003; Fletcher & Shaw, 2011; Johnson, 2011; MacLean, Meyer, & Estable, 2004; Matheson, 2007; Tilley, 2003). The SRS is also known as Voice Recognition Software (VRS) and the terms have been

used interchangeably in journals. There are several SRS available such as Dragon Naturally Speaking, MacSpeech Scribe, Braina, Sonic Extracter or Via Voice. Dragon Naturally Speaking appears to be the most preferred SRS by researchers.

What needs to be pointed out is that SRS are "speaker dependent" and do not automatically transcribe interview data. SRS only accurately transcribe the voice of a single speaker and the software needs to be trained to recognize this voice (Alcock & Iphofen, 2007). Thus, when using it for the purpose of transcribing interviews, the researcher needs to listen to the interview recording and repeats what is said. SRSs can turn voice into text as quickly as you can speak a word. Thus, there is no need to constantly rewind the audio while you try to type out the corresponding text.

Even though SRS can speed up transcription time considerably (Fletcher & Shaw, 2011); there are still issues related to accuracy (Alcock & Iphofen, 2007; MacLean et al., 2004). The software often has problems with homophone words ("one" and "won," "sun" and "son," "here" and "hear," and so on). Hence, it is important for researchers to double-check the transcription for accuracy (MacLean et al., 2004). There are many contradicting opinions regarding the pros and cons of SRS in producing transcripts (Al-Aynati & Chorneyko, 2003; Alcock & Iphofen, 2007; Johnson, 2011; MacLean et al., 2004; Park & Zeanah, 2005). Summarizing the various points of views, a majority of scholars seem to agree that it speeds up the transcription time and reduces stress.

Albeit above we emphasized the importance of preparing transcripts, we do not see the use of SRS as a threat to the qualitative inquiry process: As the software cannot handle the process fully automated, the researchers is still immersed in the preparation of the transcript by listening to the recording and repeating whatever was said. Thereby, the researcher still maintains closeness to the data. Furthermore, if it is desired to associate the media file with the transcript, the researcher needs to go through the recording and transcript again to set time stamps. While doing this, the accuracy of the transcript can be double-checked and if necessary, corrections can be made. In addition, emerging thoughts, ideas and first insights can be jotted down in memos.

Looking to the future, developers and researchers are working hard to improve various aspects of SRS, such as the support for differences in speaking pattern (Kwon, Kim, & Choeh, 2016), real time correction of errors (Wald, Bell, Boulain, Doody, & Gerrard, 2008), emotion recognition (Chenchah & Lachiri, 2015; Han, Li, Ruan, & Ma, 2014; Mariooryad & Busso, 2014), social intelligence, translation cloud (Vicsi, 2012), and many more. Hence, it is safe to assume that artificial intelligent technology will not be long to vastly expand the current capabilities.

With the advancement of speech recognition technology, we cannot underestimate its potential to replace current human transcription at all. The "deep learning" technology, which developers currently working on, will allow the computer to identify the context of conversation, which eventually will also solve the homophone problems (Tuttle, 2015). This definitely will increase the accuracy of transcriptions.

Meanwhile, the development in emotion recognition will allow computers to recognize human emotion, which will also be very useful for transcription. It is therefore very feasible that fully automated transcription software with good accuracy and artificial intelligence will become available in the near future— replacing the human transcriptionist. Maybe soon researchers just have to play the audio recordings and the computer will automatically transcribe their data for them.

If this happens, how can researchers stay close to the data? Will they look at the transcript in a different way? How will this affect the inquiry process in qualitative research? If the computer takes over the task of transcribing, how will this effect the analytic process that is supposed to take place? And how could this influence the research findings? Will it affect the credibility of the research?

We believe that every new technology will affect the nature of knowledge discovery. If computers replace the human transcriptionist in the future, qualitative researchers need to reflect on those questions. No matter how sophisticated computer software becomes in assisting qualitative studies, researchers must always remember that qualitative inquiry is laden with in-depth understandings of the subject (Bailey, 2008; Creswell, 2012; Denzin & Lincoln, 2008). In order to develop reliable analyses interpretations, the researcher is the one who must get immersed into the data in order to fully understand the topic at hand (Evers, 2011).

Any software, whether it was developed to support the transcription or the analysis process, must be regarded as an "aiding" tool or an accelerator; but it can never fully take over the researcher's role. There are aspects in conversations or observations that only the researcher who collected the data knows about, such as slang, involuntary vocalization, response tokens, and non-verbal communication (Oliver et al., 2005). All of these aspects have meanings, provide cues, and trigger emotion within their own contexts. The interpretation of qualitative data needs human eyes and is an interpretative act (Evers, 2011). Even if supported by technology, we strongly suggest that only the researcher himself/herself should be one to finalize and shape the final text. In the end, the researcher is the one who is responsible for the overall quality of the project.

Conclusion

Transcription is an important process and the first analytic step in working with qualitative data. The decision on what and how to transcribe should be decided at an early stage of the study and is dependent on the objectives of the project, the chosen methodological approach, and the research context. Researchers can select from among a number of transcription tools—including CAQDAS such as ATLAS.ti— that, in addition to facilitating the transcription process, can also be used for the continuing analysis. While in the future speech recognition software may eradicate the need for typing, researchers still have to invest their time in listening, dictating viewing, and checking the media file while it is being transformed into a transcript. In this paper we have highlighted the many benefits of getting involved in the

preparation of the transcript oneself. It actuates the analytic process as it engages researchers closely with the data, which in turn facilitates the process of developing an in-depth understanding over time and to draw credible conclusions.

References

Al-Aynati, M. M., & Chorneyko, K. A. (2003). Comparison of voice-automated transcription and human transcription in generating pathology reports. *Archives of Pathology & Laboratory Medicine, 127*(6), 721-725.

Alcock, J., & Iphofen, R. (2007). Computer-assisted software transcription of qualitative interviews. [Article]. *Nurse Researcher, 15*(1), 16+.

Bailey, J. (2008). First steps in qualitative data analysis: transcribing. *Family Practice, 25*(2), 127-131. doi: 10.1093/fampra/cmn003

Chenchah, F., & Lachiri, Z. (2015). Speech emotion recognition using auditory features. *International Journal of Applied Engineering Research, 10*(23), 43412-43415.

Creswell, J. W. (2012). *Educational Research : Planning, Conducting and Evaluating Quantitative and Qualitative Research* (4th Ed ed.). Boston, MA: Pearson Education.

Creswell, J. W. (2013). *Qualitative Inquiry And Research Design : Choosing Among Five Approaches* (3rd ed.). USA: SAGE Publication.

Demster, P. G., & Woods, D. K. (2011). The Economic Crisis Through the Eyes of Transana. *Forum Qualitative Sozialforschung / Forum: Qualitative, 12*(1).

Denzin, N., & Lincoln, Y. (2008). Introduction : The Discipline and Practice of Qualitative Research. In N. Denzin & Y. Lincoln (Eds.), *Collecting and Interpreting Qualitative Materials* (pp. 1-44). USA: Sage Publications, Inc.

Easton, K. L., McComish, J. F., & Greenberg, R. (2000). Avoiding Common Pitfalls in Qualitative Data Collection and Transcription. *Qualitative Health Research, 10*(5), 703-707. doi: 10.1177/104973200129118651

Evers, J. C. (2011). From the Past into the Future. How Technological Developments Change Our Ways of Data Collection, Transcription and Analysis. *Forum : Qualitative Social Research, 12*(1).

Fletcher, A. K., & Shaw, G. (2011). How voice-recognition software presents a useful transcription tool for qualitative and mixed methods researchers. [Report]. *International Journal of Multiple Research Approaches, 5*(2), 200+.

Hammersley, M. (2010). Reproducing or constructing? Some questions about transcription in social research. *Qualitative Research, 10*(5), 553-569. doi: 10.1177/1468794110375230

Heath, C., Hindmarsh, J., & Luff, P. (2010). *Video in Qualitative Research : Analysing Social Interaction in Everyday Life*. London: SAGE Publications Inc.

Johnson, B. E. (2011). The speed and accuracy of voice recognition software-assisted transcription versus the listen-and-type method: a research note. *Qualitative Research, 11*(1), 91-97. doi: 10.1177/1468794110385966

Kiyimba, N., & O'Reilly, M. (2015). The risk of secondary traumatic stress in the qualitative transcription process: a research note. *Qualitative Research*. doi: 10.1177/1468794115577013

Kwon, S., Kim, S.-J., & Choeh, J. Y. (2016). Preprocessing for elderly speech recognition of smart devices. *Computer Speech & Language, 36*, 110-121. doi: http://dx.doi.org/10.1016/j.csl.2015.09.002

Lapadat, J. C. (2000). Problematizing transcription: Purpose, paradigm and quality. *International Journal of Social Research Methodology, 3*(3), 203-219. doi: 10.1080/13645570050083698

Lapadat, J. C., & Lindsay, A. C. (1999). Transcription in Research and Practice: From Standardization of Technique to Interpretive Positionings. *Qualitative Inquiry, 5*(1), 64-86. doi: 10.1177/107780049900500104

MacLean, L. M., Meyer, M., & Estable, A. (2004). Improving Accuracy of Transcripts in Qualitative Research. *Qualitative Health Research, 14*(1), 113-123. doi: 10.1177/1049732303259804

Mariooryad, S., & Busso, C. (2014). Compensating for speaker or lexical variabilities in speech for emotion recognition. *Speech Communication, 57*, 1-12. doi: http://dx.doi.org/10.1016/j.specom.2013.07.011

Matheson, J. L. (2007). The voice transcription technique: use of voice recognition software to transcribe digital interview data in qualitative research. [Report]. *The Qualitative Report, 12*(4), 547+.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative Data Analysis : A Methods Sourcebook* (3rd ed.). USA: SAGE Publications, Inc.

Mishler, e.g., (2003). Representing Discourse : The Rhetoric of Transcription. In Y. Lincoln & N. Denzin (Eds.), *Turning Points in Qualitative Research : Tying Knots in Handkerchief* (Vol. 3, pp. 297-326). Wlnut Creek, CA: Altamira Press.

Oliver, D. G., Serovich, J. M., & Mason, T. L. (2005). Constraints and Opportunities with Interview Transcription: Towards Reflection in Qualitative Research. *Social Forces, 84*(2), 1273-1289.

Park, J., & Zeanah, A. E. (2005). An evaluation of voice recognition software for use in interview-based research: a research note. *Qualitative Research, 5*(2), 245-251. doi: 10.1177/1468794105050837

Paulus, T., Lester, J. N., & Dempster, P. (2014). *Digital Tools for Qualitative Research*. London: SAGE Publication Inc.

Skukauskaite, A. (2012). Transparency in Transcribing: Making Visible Theoretical Bases Impacting Knowledge Construction from Open-Ended Interview Records. *Forum : Qualitative Social Research, 13*(1).

Tilley, S. A. (2003). "Challenging" Research Practices: Turning a Critical Lens on the Work of Transcription. *Qualitative Inquiry, 9*(5), 750-773. doi: 10.1177/1077800403255296

Tilley, S. A., & Powick, K. D. (2002). Distanced Data: Transcribing Other People's Research Tapes. *Canadian Journal of Education, 27*(2/3), 291-310.

Tuttle, T. (2015). The Future Of Voice-Activated AI Sounds Awesome. (Generic). Retrieved from techcrunch.com website: http://techcrunch.com/2015/03/06/the-future-of-voice-activated-ai-sounds-awesome/

Vicsi, K. (2012). *Thinking about the present and future of the complex speech recognition*. Cognitive Infocommunications (CogInfoCom), 2012 IEEE 3rd International Conference, 371-376.

Wald, M., Bell, J.-M., Boulain, P., Doody, K., & Gerrard, J. (2008). Correcting automatic speech recognition captioning errors in real time. [journal article]. *International Journal of Speech Technology, 10*(1), 1-15. doi: 10.1007/s10772-008-9014-4

Wilson, R. (2002). Research notes. *Nursing Standard, 17*(5), 24.

Wright, M., & Russell, A. (2010). Using NVivo Audio-Coding : Practical, Sensorial and Epistemological Considerations. *Social Research Update, Autumn 2010*(60). Retrieved from http://sru.soc.surrey.ac.uk/SRU60.pdf

Acknowledgments

About The Authors

Rohaiza Zakaria

PhD candidate, Dept. of Technical and Engineering Education, Faculty of Education, Universiti Teknologi Malaysia. Email : ixa.zakaria@gmail.com

Aede Hatib Musta'amal

Senior Lecturer, Dept. of Technical and Engineering Education, Faculty of Education, Universiti Teknologi Malaysia.

Nor Fadila Mohd Amin

Senior Lecturer, Dept. of Technical and Engineering Education, Faculty of Education, Universiti Teknologi Malaysia.

Halmi Mashindra Saleh

Lecturer, Dept. of Design and Visual Communication, Politeknik Ibrahim Sultan, Johor, Malaysia

Article Information