Transcribing in ESL: An Investigation into Role, Accuracy, and Skills

John G. Patkin
The Education University of Hong kong, patkin@eduhk.hk

Transcribing in ESL: An Investigation into Role, Accuracy, and Skills

Abstract

The role, accuracy, and skills of a team of ESL researchers who transcribed more than 3,000 English language learning histories (LLH) from university students in Hong Kong is examined in this paper. The paper provides an insight into the role of transcribers, how they approach their work, the problems they face, and how they overcome them including the conflict of their prior English language learning. A self-administered semi-structured interview and thematic analysis were used in this qualitative study. The findings show that transcribers learned experientially as they combined project guidelines and prior experience to achieve accuracy Transcribers feel more comfortable working on data that is closer to their English as a second language (ESL) background; however, this also contributed to personal conflict such as correcting grammatical errors. The outcomes suggest training, clear guidelines from supervisors, and incorporating feedback from transcribers can improve the richness and accuracy of data which is of great importance to second language data collection.

Keywords

transcription, ESL, interviews, thematic analysis

Creative Commons License

Acknowledgements

Transcribing in ESL:
An Investigation into Role, Accuracy, and Skills

John G. Patkin
The Education University of Hong Kong

The role, accuracy, and skills of a team of ESL researchers who transcribed more than 3,000 English language learning histories (LLH) from university students in Hong Kong is examined in this paper. The paper provides an insight into the role of transcribers, how they approach their work, the problems they face, and how they overcome them including the conflict of their prior English language learning. A self-administered semi-structured interview and thematic analysis were used in this qualitative study. The findings show that transcribers learned experientially as they combined project guidelines and prior experience to achieve accuracy Transcribers feel more comfortable working on data that is closer to their English as a second language (ESL) background; however, this also contributed to personal conflict such as correcting grammatical errors. The outcomes suggest training, clear guidelines from supervisors, and incorporating feedback from transcribers can improve the richness and accuracy of data which is of great importance to second language data collection.

Keywords: transcription, ESL, interviews, thematic analysis

Introduction

The Hong Kong Archive of Language Learning (HALL) is a web-based archive of more than 3,000 audio interviews about English language learning experiences (HALL, 2016). The archive aims to provide support and inspiration for English language learners and instructors, while the interviews and corresponding transcripts are being analyzed to better inform practice. This paper focuses on the transcription of the interviews with multilingual undergraduates in Hong Kong universities. A broader discussion of the project was published earlier (see Lee & Patkin, 2017).

The review of literature will discuss how transcription is employed in academic research. It will lead to three research questions that focus on the role of transcribers, their thoughts on accuracy, and the skills used in the preparation of transcripts. Through the answers of a semi-structured self-administered e-interview, the findings will reveal that transcribers share a belief in the need to act as a gatekeeper but are challenged by previous learning that conflicts with the need to provide a genuine transcript. The study concludes with a recommendation that the time and resources allocated to the process of transcription should be emphasized more at all stages of the research process and ESL transcribers should be employed for their insights into their familiar variety of English.

Review of Literature

Transcription is commonly used to provide a textual representation of LLHs captured during field interviews to provide researchers a deeper understanding of personal experience while expanding methods of investigation (Barkhuizen et al., 2013). Transcription is widely

employed in other areas of academia, law, and health care (Tong et al., 2007) yet calls for a closer examination of its methods have been made over the years (Bucholtz, 2000; Halcomb & Davidson, 2006; Poland, 1995; Riessman, 2002), particularly because transcripts are used to make life-changing decisions (O'Connell & Kowal, 2008) in fields such as law and medicine. Others have cautioned novice researchers on the use of transcription, suggesting it be employed carefully and alongside other data sources (Cohen et al., 2013). While the body of literature on the methodological role of transcription (see Bird, 2005; Breiteneder et al., 2006; Edwards, 2014; Jenks, 2012; Lapadat, 2000; MacWhinney, 2014; Mondada, 2013; Poland, 1995; Powers, 2005; Tilley, 2003) and the challenges of transcribing nonnative varieties such as dialects are growing (Bucholtz, 2000; Lampert & Ervin-Tripp, 2014), there is a noticeable gap in empirical studies on the role of ESL transcribers.

What is Transcription?

Transcription is an orthographic representation and interpretation of sounds and activities (Bird, 2005; Edwards, 2014; Halcomb & Davidson, 2006) of humans and their interactions (Roberts, 1997) and is strongly influenced by the operator (Mondada, 2013) i.e., the transcriber as an "agent of change" (Lindsay & O'Connell, 1995, p. 102). The outcome of transcription is a customized (Mondada, 2013) and lasting "manipulable" product that can be "quoted, sorted, copied, and inspected" (Lapadat, 2000, p. 204). Extremes, though not direct opposites, dominate the definition of transcription - from the all-encompassing transcript (open) to summaries and selected sections transcribed with custom conventions (closed), all of which come in a variety of characters, some in color, that attempt to represent aspects of a speech event beyond words. An example of this with mark-ups in the VoiceScribe (VS) software, which was used in HALL, is illustrated below (Figure 1).

Figure 1
Open and Closed Transcription in VoiceScribe
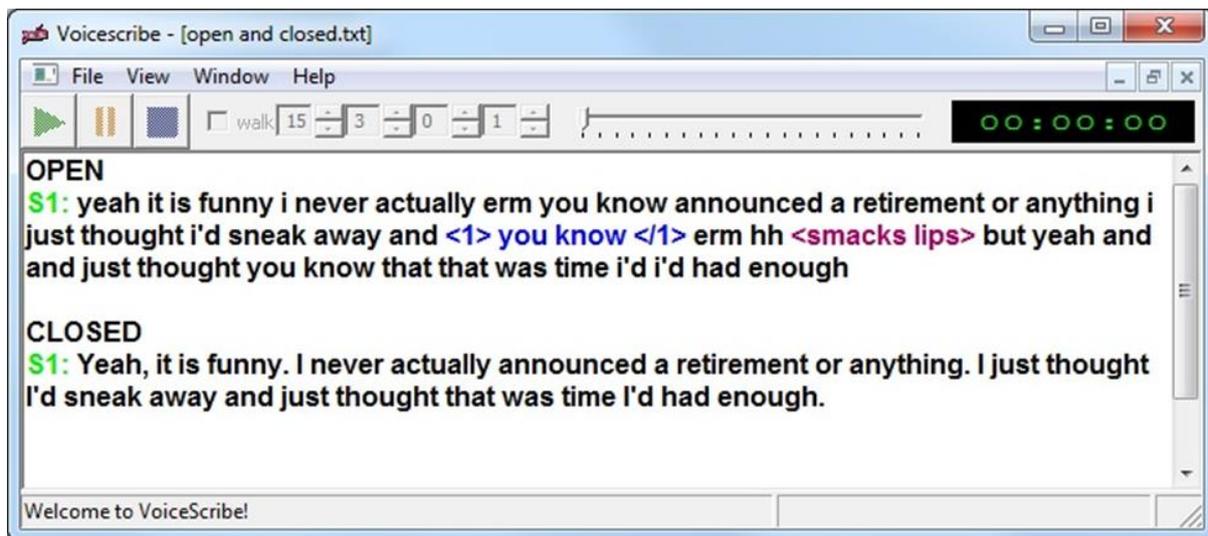
The terms verbatim (Poland, 1995), non-standard, open (Jenks, 2012), and natural (Oliver et al., 2005) are used in varying disciplines to describe a transcript that aims to capture all sounds and utterances in a recorded act of human communication. These "encompassing transcripts" attempt to provide an unfiltered textual representation of every noise, breath, pause, stutter, inflection, pronunciation, and grammatical deviation. Textual symbols and color illustrate a variation of utterances, while field notes provide background. Selective transcripts

focus on the data needs of specific projects (Jenks, 2012; O'Connell & Kowal, 2008; Oliver et al., 2006). Speech pathologists may be interested in pauses, breaths, and stuttering; hesitation and repeated utterances may inform psychologists; linguists may be interested in coinage as part of an investigation into language shift. HALL transcripts have been used to investigate the role of literature in English language learning (Lee & Patkin, 2016).

Transcriptions cannot be used as a standalone record of an event and it is impossible to capture non-verbal live action that permeates the senses of interlocutors and their observers (Jenks, 2012; MacWhinney, 2014; Powers, 2005). Constantly reviewing interviews is time-consuming and expensive, so the transcript acts as a substitute. Like drama scripts, transcripts are interpretations with limitations (Jenks, 2012; Poland, 1995; Riessman, 2002). Transcripts are the outcome of an incomplete negotiation of what is heard and valued by the transcriber. HALL researchers were encouraged to transcribe their own field interviews in order to provide context to strengthen interpretations.

Planning Transcription

The role of transcription is often overlooked in the preparation of research proposals and reports. Proposals can underestimate the time and costs, while publications provide scant detail on the methods, techniques, conventions, resources, human interaction, and technical details of transcribing. Despite a disparity among scholars in the techniques of transcription, there is wide agreement on the need to plan (Bird, 2005; Breiteneder et al., 2006; Halcomb & Davidson, 2006; Jenks, 2012; MacWhinney, 2014; McLellan et al., 2003). One of the worst outcomes would be to modify transcription goals after starting due to an underestimation of the required resources (Tilley, 2003). Planning HALL's transcription was a combination of published guidelines and experience from prior projects.

Transcription Conventions

Employing a uniform system of conventions ensures large data sets from a variety of projects can be compared (Edwards, 2014; Mergenthaler & Stinson, 1992). Apart from linguistic markers, conventions can be used to establish uniformity in spelling (Breiteneder et al., 2006). Project transcribers need to agree on how best to represent hundreds of conversations in a uniform way (Jenks, 2012); however, they should be wary of uniformly attempting to map a set of conventions from one project to another (Lapadat, 2000; McLellan et al., 2003). HALL used a modified version of conventions from the Vienna Oxford International Corpus of English (Appendix 1; VOICE, 2013), and British spelling was adopted.

The procedure for preparing interviews for transcription included editing the beginning and end to eliminate doubt about starting points, length, and content. When multiplied by more than 6,000, a figure that included a first and second transcriber, hundreds of hours of labor were saved by transcribing only the answers rather than the interviewers' opening question and concluding remarks and pleasantries[1]. The unique keyboard shortcuts in VS allows users to control audio while typing in the same programme on a PC (VOICE, 2013). VS helped transcribers check scripts when they typed, as incorrect mark-ups would not display the corresponding color-codes. The color scheme also reduced the number of symbols and decoding skills required by the human analyst.

The semi-closed protocols in HALL reduced the workload for transcribers, but they were required to log instances where utterances, pauses, and context affected the interview.

---

[1] The editing did not affect the content of the interviews. If the interviewer's question was referenced by the interviewee, it was included in the transcript.

These included repeated hesitations, pauses longer than four seconds, help given to interviewees by others to respond to a question and mechanical events that affected the recording such as loud noises, and recording failures. In the example below (Figure 2), the interviewee's hesitation is illustrated with six "erms" and several pauses.

Figure 2
Mark-ups Illustrating the Use of Fillers and Pauses
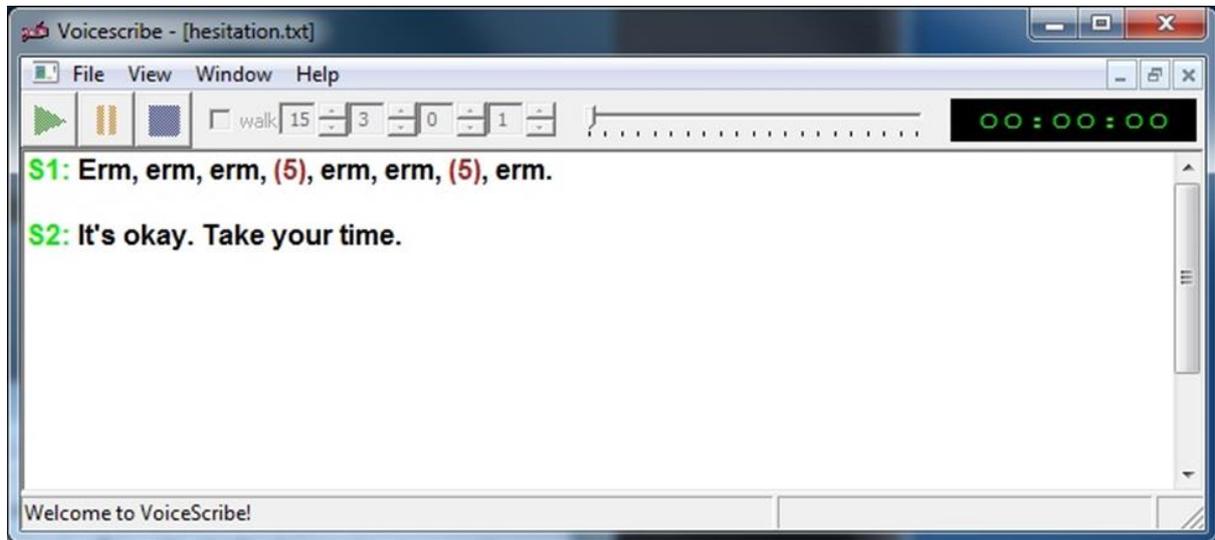
Transcriber Training

Investigators need to strike a balance between resources and research goals when choosing how transcription will be completed. Project transcription by the investigators can be avoided through outsourcing, but if team researchers do it, they will have a better feeling for the data (Bird, 2005; Halcomb & Davidson, 2006; Lapadat, 2000; O'Connell & Kowal, 2008; Tilley, 2003) and the process is appealing to people interested in human interaction beyond research objectives (Powers, 2005). It is common practice for researchers to employ transcription companies to save time, but it can be more expensive, and some level of time investment is still required for training and checking outsourced work (Jenks, 2012; Mergenthaler & Stinson, 1992). Employing student helpers on the premise they may be interested in the research may be counterproductive as their lack of experience may produce poor transcripts (Powers, 2005). Bird's (2005) reflection and Matheson's (2007) overview suggest positively framing the task of transcribing despite preconceived negativity. This helps trainee transcribers make their own decisions about whether it is an enjoyable exercise, but they should be warned that the process can be time-consuming (Bird, 2005; Lampert & Ervin-Tripp, 2014), and once trained, good transcribers should be retained (Powers, 2005). As the lead field researcher and chief transcriber, I leveraged my experience to emphasize the benefits of transcribing were more than financial and that being part of the project would strengthen their understanding of the research process. In workshops, I used examples from previous projects and explained how the training would give them a skill they can use for future studies and jobs.

Transcription systems need to be well designed and thoroughly explained (Edwards, 2014; McLellan et al., 2003; Riessman, 2002) as novices find it difficult to use a familiar language with non-standard forms (O'Connell & Kowal, 2008). If the transcriber is only working on the recordings, then the researcher needs to provide more background about the

aims, purposes, and processes, including the description used in publications, to facilitate a better understanding (Lapadat, 2000). In-house, outsourced, or researcher-transcribed, training ensures analysts will be able to read a truer account of what was said rather than what a transcriber thinks she heard (Breiteneder et al., 2006; Bucholtz, 2000; Poland, 1995). To ensure rigor, training should be ongoing and reflexive; diaries, meetings, and discussions should be used alongside transcript reviews and feedback (Lapadat, 2000).

HALL transcripts were checked against original recordings (McLellan et al., 2003) and if a high frequency of errors such as standardization and lack of conventions were found, the transcriber was asked to revise and resubmit. Standardization included the editing and correction of an interlocutor's speech and missed conventions included pauses, fillers, context, unintelligible speech, coinages, other languages, and laughter (see Appendix 1). This iterative process helped established a benchmark but was handled delicately to avoid demoralizing team members. Transcribers increased their confidence once they no longer had scripts returned, and we explained that we trusted their submissions because fewer corrections were made during the second round of checking. Guidelines and conventions ensured uniformity and authenticity. If the transcribers had standardized the speech, the transcript would not reflect the oral competency of the interlocutors or how they communicated with each other, particularly if there had been misunderstandings. The conventions also provided insight into the meaning of some responses such as pauses, filler, and laughter to mask embarrassment (Walkinshaw & Kirkpatrick, 2014), coinages, and code-switching which highlighted the relationship between the use of English and a first language (VOICE, 2013).

Errors attributed to a lack of subject knowledge and human ability (Reason, 1990) or even fatigue (Poland, 1995) can affect the transfer of audio to text. Without guidance, transcribers are inclined to insert their interpretations (Bird, 2005; Edwards, 2014; Tilley, 2003). Common errors and mistakes are attributed to "additions and omissions" (Pavlenko, 2007, p. 173). The omission of clarifying words such as "no" (Mergenthaler & Stinson, 1992) or using a homophone when mishearing (Lindsay & O'Connell, 1995) can have a major effect on research.

Error-ridden transcripts were prepared for discussion during workshops. When asked how the errors occurred, colleagues drew on their own experiences. Comparing work with colleagues helps to understand common goals and eliminate uncertainty (Powers, 2005) and it is easier to follow the development of transcripts if there is a verbatim text (Halcomb & Davidson, 2006). Through training and practice, transcribers become more adept with software and machinery (Allwood et al., 2000; Bird, 2005) which allows them to concentrate more on the accuracy of the content rather than the mechanism for processing and developing it.

Challenges in Transcription

No matter how much training project leaders provide, transcribers still make simple mistakes based on personal interpretations (Roulston et al., 2003) including the desire to correct non-standard grammar (MacWhinney, 2014), an area that has received little attention (Roberts, 1997). Some researchers suggest grammar correction is acceptable (Barkhuizen et al., 2013) and unavoidable due to personal bias (Lindsay & O'Connell, 1995; Oliver et al., 2006; Powers, 2005); however, the HALL conventions dictated that the original would be used which challenged the ESL transcribers.

Early transcripts in HALL revealed one colleague's belief that non-standard forms should be corrected as the original utterance made the interviewee look bad. Poland described such instances as a "tidy up" (1995, p. 296); Lindsay and O'Connell's study (1995) found "adverbs and conjunctions as the most frequent additions" (p. 111). Familiarity with varieties of English and general knowledge also affected transcripts as some of the researchers were

unfamiliar with the topic mentioned by the interviewee, especially places outside Hong Kong where they had traveled. Transcription is also affected by the quality of audio recordings (Tilley, 2003) which can lead to misunderstandings.

Unintelligible utterances were handled in two stages. They first required the transcriber to revisit the utterance after completing a first draft because vocabulary was often repeated and sometimes easier to identify in another collocation. If it remained unintelligible, a note was typed into the file header so that the second transcriber could check it (Poland, 1995). Despite our best efforts, some utterances remained unknown. Such cases are considered as limitations of transcription (O'Connell & Kowal, 2008). Interviews were conducted in English, however, some interviewees reverted to their first language (L1) when accessing thoughts. To maintain their original meaning (Pavlenko, 2007), these dialogues were presented in their original form along with a translation.

Transcription is criticized for its inability to capture color and smell along with the feelings of the interlocutors (Lapadat & Lindsay, 1999). Other stages in the research process such as data collection, processing, and analyses also have their flaws. Transcribing is more expensive than some realize due to the many hidden hours that are required to decipher different accents and poor recordings. Such costs can increase if a transcriber misinterprets instructions resulting in hundreds of hours of work that are meaningless to an investigation. For example, if a transcriber corrects a speaker's grammar it might jeopardize an investigation into tense markers. It is therefore prudent and cost effective to check a small selection of drafts to ensure the transcriber is following instructions and correct any problems early on (Poland, 1995).

Research Questions

The review of literature has shown us there is a rich body of information that can be used to guide transcribers however there is a lack of focus on how ESL transcribers interact with conventions. These gaps in the literature and the growing role of ESL in language research led to an examination centering on the following three questions:

RQ1: How do ESL transcribers define their role?
RQ2: How does insight contribute to an ESL transcriber's accuracy?
RQ3: How do ESL transcribers apply their skills?

The research questions aim to contribute to the body of knowledge on transcribing with a focus on the growing area of ESL researchers. By answering these questions, investigators will be better informed about the needs of ESL transcribers while practitioners will benefit from the shared experiences of peers. The methods and outcomes of the investigation will be discussed in the following sections.

Methodology

A self-administered semi-structured interview and thematic analysis were used for this qualitative study investigating the role, accuracy, and skills of ESL transcribers. All 13 of the project's transcribers were approached to take part but four declined. All responses have been anonymized[2].

---

[2] The present study fell under the auspices of the Digital Archive of English Language Learning Histories, which was approved by the Human Subjects Ethics Committee at City University of Hong Kong.

Participants

All participants in the present study had attended at least one transcription training workshop. The workshops were followed up with e-coaching which included practice transcription and ongoing feedback. The workshops introduced the nature of the project, the data collection process, and file management. The training aimed to familiarize the researchers with the transcription conventions (Appendix 1) employed for the project, identify common themes and language, and forewarn them about common errors and mistakes. The profile of each respondent is tabled in Figure 3 which shows their employee role, level of academic achievement, languages spoken, gender, and age.

Figure 3
Profile of Respondents

Respondent	Role	Academic status	Languages spoken	Gender	Age
T1	TR, PW	B, M.	English, Cantonese (L1)	F	>40
T2	TR, IV	B, M	English, Cantonese (L1)	F	35-40
T3	TR, IV, PW	B, M, P (IP)	English, Cantonese (L1)	F	30-40
T4	TR, IV, PW	B.	English (L1), Cantonese	M	20-25
T5	TR, IV, PW	B.	English, Cantonese (L1)	F	20-25
T6	TR, IV, PW	B.	English, Cantonese (L1)	F	20-25
T7	TR, IV, PW	B. M.	English, Cantonese (L1)	M	25-30
T8	TR, IV, PW	B (IP)	English, Cantonese (L1)	F	20-25
T9	TR, IV, PW	B. M.	English, Cantonese (L1)	F	25-30

Note. TR=transcriber; PW=project worker; IV=Interviewer; IP=In progress; B=Bachelor; M=Master; P=PhD.

All respondents had transcribed interviews and proofread colleagues' work while all but one had conducted primary interviews. Most had completed an undergraduate degree, and more than half had a postgraduate qualification. All had studied in English at university level and spoke Cantonese as an L1 except T4, who was a native English speaker.

Data Collection

The self-administered seven-item interview was emailed to respondents who were asked to type their replies. The e-interview format was chosen as much of the communication for the project had been conducted via email. It also allowed the respondents to curate their answers in their own time and without prompting (Burns, 2010; Ratislavova & Ratislav, 2014). As I had acted as their leader in the field, I was wary of a possible power distance effect (Hubbell, 2003) and therefore felt it was important for them not to feel as if they had to respond or be "supervised" while answering.

The first six questions addressed four main issues including the role of a transcriber, prior knowledge of the field, problems, and solutions, and what was learned. The final question asked for additional comments that may have led to topics that were not already addressed (Castillo-Montoya, 2016). The table below (Figure 4) aligns the six open questions with the RQs. To ensure the RQs were answered, specific questions were asked; however, some responses provided greater insight into the overall study.

Figure 4
RQs Aligned to Interview Items

RQ	Question
1	How would you describe your role as a HALL transcriber?
	What did you know about transcription before you started working on HALL?
2	What was the difference between transcribing for HALL and what you already knew about transcribing?
	What kinds of challenges did you face during transcription?
3	What did you learn while transcribing?
	How did you meet those challenges?

Data Analysis

An iterative process was used for the data analysis. Thematic analysis (Aronson, 1995) was employed manually and with the assistance of software (Elliott, 2018). The first stage was the reading of all responses, the second was the generation of keywords through a software programme, and the third was thematically linking the keywords and RQs to the responses of the interviewees. The manual and electronic sorting of the data ensured its rigor as the final stage was linked to the interviewee's original statements.

The reading stage allowed me to understand how the interviewees had responded to the questions and whether there was a need to clarify any of their statements. All nine transcribers provided responses to the six questions in the self-administered interview. Only one (T7) responded to the seventh question with a comment stating he was surprised how challenging transcribing can be. The responses were then collated into one file and WordSmith Tools (Scott, 2008) was employed to create a list of keywords. The total word count was 3,366. After conjunctions, prepositions, articles, and unassociated words were eliminated, the word list was conflated to 372. The words were grouped according to RQ keywords – "Role," "Accuracy" and "ESL."

Using WordSmith's concordance feature (Johnson & Ensslin, 2006), words were listed as frequencies. For example, know, knowing, knew, and knowledge (n=13) and learn, learned, learning, learnt, and lessons (n=18) were conflated to the root word "experience" which featured throughout Bird's (2005) reflection of being transcriber. In RQ1, "Role" was linked to "experience" and "knowledge" through T1's statement: "My first experience of transcribing occurred in working on HALL. There was not much to be compared with my previous knowledge on transcription."

Figure 5
Word List Aligned to RQs

RQ	RQ key words	Root words
1	Role	self, experience, ownership, task
2	Accuracy	accuracy, sound, rules
3	ESL	understand, culture, language, grammar, difference

Powers (2005) associates accuracy with the sounds of interviewees. Sound, speak, speaker, speakers, speaking, speaks, speech, speeches, spoke, spoken, talk, and talking (n=30) were conflated to sound. In RQ2, accuracy was linked to sound. T2's statement in response to the question about the challenges faced during transcription provides an example of this link:

> When the interviewee has a strong regional accent, he/she speaks too fast or at extreme pitches either too high or too low, doesn't fully pronounce the words, frequently chimes in in the middle of others' responses and the background of recording is full of noise.

The third RQ centered on the theme of ESL. Words such as differ, different, and differently (n=14) highlighted the conflict faced by the transcribers as they tried to adapt to other accents when typing up interviews. The relationship to ESL is linked to the problems faced by non-native speakers (Lampert & Ervin-Tripp, 2014) or researchers working with regional dialects and accents (Bucholtz, 2000). Such experiences include those of ESL transcribers working with differences in culture, grammar, and language (Breiteneder et al., 2006). T4 explains he was unfamiliar with some of the interviewee's accent: "It was a little difficult at the beginning, but I eventually got accustomed to it and was able to understand different accents more easily."

Findings

The findings are presented in three sections in order of the RQs. Firstly, there are results on how ESL transcribers define their role. Secondly, the results of the techniques ESL transcribers use to ensure accuracy are presented. Finally, the results of the third RQ on how ESL transcribers apply their skills are presented.

RQ1: How do ESL Transcribers Define Their Role?

HALL transcribers defined their role as an evolving gatekeeper that developed skills organically. As they reflected on the rules and conventions of transcribing for HALL, the respondents drew on their past experiences and how they transformed their understanding of the importance of their role in the research process. They knew they had to follow a system of conventions in the transformation of audio to text.

In excerpt 1, T8 was one of three honors students who had joined the HALL research team as part of a final year project. She recognized the importance of her gatekeeping role and sought to clarify the rules of the project and the role of a transcriber with a friend.

Excerpt 1

> We also have to follow guidelines, e.g., mark-up conventions, when we transcribe audio records. I have a friend who also does transcriptions part-time, but I never heard of these from her. (T8)

In excerpt 1, T8 compared her role with a friend while T4 (Excerpt 2) and T6 (Excerpt 3) grew to understand the importance of their contribution to the project.

Excerpt 2

> Since my only experience with transcription prior to HALL was during my time working as an English Broadcast Journalist, I saw it as an informal process which simply allowed people to locate information, they might need quickly for purposes such as quoting. Transcribing for HALL was a completely new experience for me as the whole process was much more structured and professional compared to my previous experiences. (T4)

Excerpt 3

> I have done some transcription before for a different project. It was a completely different experience as there weren't any guidelines or rules; I was simply told to type up the dialogues on a word file from beginning to end. There were no symbols, no revisions, and certainly no VoiceScribe. I would describe it as a very amateur experience. (T6)

The experience changed T4 and T6's perception of a transcriber from an amateur role to a professional one. That professional role was defined by T7 in excerpt 4 who felt he was a conduit between interviews and investigators.

Excerpt 4

> I have to transform experiences which students shared orally into written texts, which researchers can use in their studies. (T7)

T7 defined his role as a stage in the research process. The limitations of this role are acknowledged by T9 (Excerpt 5) who also accepted the role of gatekeeper.

Excerpt 5

> Although it was understandable that the transcription could only be a representation and could not capture every detail in the time of interview, I tried to identify the speech as accurate as I could without changing their use of words and grammar. (T9)

The transcribing guidelines surprised T8, T4, and T6 who acknowledged transcription requires a variety of known and emerging skills and resources to meet the goals and requirements, T7 defined the transcriber's role in the research process while T9 revealed the limitations.

RQ2: How Does Insight Contribute to an ESL Transcriber's Accuracy?

Transcribers rely on existing skills and utilize a variety of tools including audio equipment and the Internet to ensure accuracy.

T5 noted the importance of listening carefully (Excerpt 6). She was wary of losing concentration and omitting utterances that are critical for research.

Excerpt 6

> ….without listening actively and consciously to what the interviewees are saying, it's easy to miss noting an expression or the tone of the speaker, which are critical to the analysis process in the research. It is essential to be aware of this because it will prevent the original meaning of a transcript being altered. (T5)

Omitting an utterance is similar to leaving out unintelligible speech. Returning to unintelligible speech after completing the first draft allowed T6 (Excerpt 7) to maintain momentum instead of being stuck.

Excerpt 7

> And if I really can't make out what the interviewee is saying, I would return
> to it after a couple of hours or even a day – sometimes it helps so that I don't
> get stuck on the same part again and again. (T6)

Identifying unintelligible speech may improve over time. Transcribers may also consider their working environment and equipment. In excerpt 8, T9 improved her comprehension by using an over-ear headphone instead of an ear bud.

Excerpt 8

> Using better equipment (headphone instead of earphone) and working in a
> quiet environment would help enhance the accuracy of transcription. (T9)

Once the transcribers had set-up their equipment, the next stage was increasing their output while maintaining accuracy. In Excerpt 9, T1 explains that she tried to type as quickly as the interviewees.

Excerpt 9

> Typing fast and accurately while keeping up with the audio was challenging
> at first. I do type a lot on a daily basis for emails and document processing.
> However, more practice was needed to type at a pace that's demanded of
> audio transcriptions. (T1)

Sometimes transcribers could listen clearly but were unfamiliar with the language used by the interlocutors. In excerpt 10, T7 was able to fill in the gaps by using open-source websites.

Excerpt 10

> I also use Google to find titles of works and some unfamiliar words.
> Wikipedia is very useful for finding characters and things happened in the
> stories they talk about. (T7)

ESL transcribers define accuracy as critical for research outcomes. Transcribers need to focus to listen and utilize tools such as better audio equipment and the internet to improve their comprehension.

RQ3: How do ESL transcribers apply their skills?

Transcribing other ESL speakers invoked the transcribers' own learning and beliefs toward the use of English. Some of the transcribers were able to utilize their prior learning while others found it a hindrance.

As a Cantonese L1 speaker, who had been educated in Hong Kong and worked as an English teacher, T2 felt her connection to "local" interviewees improved her ability to transcribe.

Excerpt 11

> I see myself as somebody who helps translate thoughts of the interviewees as their use of English sometimes bears traits of local linguistic culture, which could be hard for non-locals to decode what they really imply. (T2)

T2 has more than 10 years' experience teaching ESL students while T4 was a fresh graduate who used English as first language. In Excerpt 12, T4's lack of experience with ESL speakers from mainland China made it harder to comprehend interviewees' accent, pronunciation, vocabulary, and proficiency in English.

Excerpt 12

> ….some had accents which I won't come across often. Therefore, accurately interpreting their words was challenging at times. ….I would not understand certain words or answers as I might lack the knowledge of certain places or education systems such as China's college entrance exams GaoKao[3]. (T4)

T4's struggle to comprehend interviewees was shared by T9 in Excerpt 13 who was confused by the pronunciation of tense markers.

Excerpt 13

> I encountered difficulties in listening to the tense markers (-ed or –t ending) during transcription. In natural conversation, those markers were not stressed and less audible. As a result, it might be difficult to decide whether the respondents had added those when they seemed to be using past tense or past participle. (T9)

The process of listening intently to other speakers of English had a positive effect T6 in Excerpt 14. She felt it improved her own speech.

Excerpt 14

> Listening to people talk made me realize the importance of clarity in speeches. It is what keeps the listener engaged in the conversation, and that's always a good thing for transcriber, because listening to a speaker struggling to find the right word to express his/her thoughts is always painful and thus makes little incentive for one to keep working. (T6)

Patience was also important for T9 in Excerpt 15. She knew that interviews generated spontaneous responses that included non-standard grammar and had difficulty transcribing language that did not follow the rules she had learned.

Excerpt 15

> ….the speech in HALL tended to be less formal, and the sentences were generally incomplete and grammatically "incorrect" as they were produced

---

[3] The GaoKao is used as a university entrance exam in the PRC.

spontaneously. Rather than following the traditional grammatical rules or sentence structures, I had to put much effort in figuring out the variation in order to convey how the respondents actually said the words. (T9)

Transcribing non-standard speech was a challenge for T9 because it did not flow as expected. In Excerpt 16, even though she knew that it was wrong to correct non-standard form, T5 initially found it difficult not to amend the interviewees' grammar.

Excerpt 16

I was particularly conscious on what the interviewees were saying, and I corrected the [note: grammatical] errors amidst listening to the conversations. I realized it was a mistake to correct them in the transcriptions, and I re-transcribe to what they were before they were corrected. It was a challenge to adapt to type exactly what the interviewees were answering in the beginning because I was often tempted to change them, even knowing it was the wrong thing to do afterwards. (T5)

Knowledge of ESL worked in two ways for HALL transcribers. Firstly, they applied their knowledge to comprehend the interlocutors. Secondly, as stated by T5 above, they experienced conflict with their prior learning as they wanted to correct non-standard speech.

Discussion

The findings have revealed how ESL transcribers define their role, how they ensure accuracy, and how they apply their skills. Some of the findings echo the views of transcribing in general while others show a direct relationship to ESL skills and experiences. In the following paragraphs, I will link the findings to the existing literature on transcribing followed by a discussion on the implications relating to ESL transcribers.

The hallmarks of a responsible and dedicated transcriber are universal.

RQ1: How do ESL Transcribers Define Their Role?

ESL transcribers play an important role in the research process. They think deeply about how their contributions affect research outcomes. They blend their existing knowledge with training, conventions, and instructions to produce scripts that are research ready. The dominant finding of RQ1 centers on ESL transcribers defining their role as gatekeepers. Their mention of rules, conventions professionalism, and the transformation of audio to text supports this postulation. Their ability to compare prior experiences and acknowledgement that they evolved while working on HALL also shows they were dedicated to their work. They also understand their role is limited to providing a textual representation of an interview.

The transcribers were able to make a distinction between the needs of HALL and other projects. In Excerpt 2, T4 stated that transcribing for HALL was "more structured and professional" while T6 felt her previous work was "very amateur" (Excerpt 3). When discussing transcribing with a friend, T8 discovered HALL was more demanding as "[she] had to follow guidelines" (Excerpt 1). These responses correlate with the literature that suggest recruits should be forewarned of the work requirements (Bird, 2005; Edwards, 2014). As training and coaching transcribers requires project funds and resources (Jenks, 2012; Mergenthaler & Stinson, 1992), the responses also support the notion of ensuring transcribers remain motivated and interested in the work (Powers, 2005). The role of transforming

interviews from audio to text which "researchers can use" (Excerpt 4) shows the value of positively framing the time-consuming experience of transcribing (Bird, 2005; Matheson, 2007). As the Chief Transcriber, I was inspired to share Powers' (2005) belief that transcribing offers an opportunity to learn from peoples' stories which I had experienced in prior projects.

RQ2: How Does insight Contribute to an ESL Transcriber's Accuracy?

Accuracy is closely connected to training and the use of tools. Training involved explaining and demonstrating the use of conventions while tools included hardware such as headphones, VS software, and the Internet. In Excerpt 6, T5 stated that she listened "actively and consciously" with the understanding that the accuracy of her work was "critical." This acknowledgement is reinforced by prior studies that also warn that misinterpretations lead to data errors (Breiteneder et al., 2006; Bucholtz, 2000; Poland, 1995) which T6 avoided by revisiting unintelligible passages on recordings (Excerpt 7). T5's responsible attitude and T6's perseverance illustrate that training and reflective practice have a positive effect on accuracy.

The relationship between the hardware and accuracy discussed above is further enmeshed when tools are combined with soft skills. T1's comment on mastering motor and cognitive skills simultaneously (Excerpt 9) and T5's remarks on concentration (Excerpt 6) echoed Jenks' (2012) and Mergenthaler and Stinson's (1992) discussion on the need for transcribers to learn the basics of listening while typing, managing headphones, and using templates. Technical errors, due to ambient noise, as experienced by T9 (Excerpt 8), can be avoided if the transcriber employs listening strategies such as headphones (Jenks, 2012; MacWhinney, 2014; Tilley, 2003). T7 illustrated the convenience of instant web resources by using "Google to find…" (Excerpt 10).

RQ3: How do ESL Transcribers Apply Their Skills?

The role of ESL transcribers emerges when they apply their language skills. Non-standard English conflicts with what they have learned, so they need to consciously detach themselves from prior learning that had mandated the correction of non-standard forms. The dilemma of transcribing non-standard grammar (O'Connell & Kowal, 2008) worked in favor of T2 who felt her familiarity with the "traits of local linguistic culture" (Excerpt 11) utilised her ESL skills. The literature warns that transcribers are inclined to add personal interpretations (Bird, 2005; Edwards, 2001; Tilley, 2003) but T2 leveraged her "local" language knowledge to "decode" the interviewees. Similarly, training helps transcribers (Edwards, 2001; McLellan et al., 2003; Riessman, 2002) and allows them to better process data as was the case with T9's approach of "figuring out the variation" (Excerpt 15), however she still faced difficulties in identifying tense markers such as "ed" which is a common trait in Hong Kong English speakers (see ACE, 2014).

Familiarity with the "local" variety of English made it comparatively easy for T2 and T9 although they had to remain vigilant in their desire to correct non-standard forms. The need for ongoing training, guidance, and checking was discovered early in the project through T5's mea culpa after she had "corrected" some interviewees' grammar (Excerpt 16). Described as a "tidy up" (Poland, 1995, p. 296), such mistakes are considered common (Roulston et al., 2003). T5's well intentioned "corrections" reflected a deeper familiarity as she had also conducted the interviews and like to chat with participants before and after recording. T5's actions align with Lapadat's (2000) suggestion that researchers can strengthen understanding with more background information. In T5's case, she used her ESL experience to provide a more detailed transcription.

The transcribers' knowledge and use of ESL worked for and against them. The positives included familiarity and comfort with interviewees of a similar linguistic background as they felt it was easier to understand what was being said. The drawback of this inside knowledge was the desire to tidy some utterances in the belief it would better represent the interviewee. This limitation was introduced in the literature review regarding the challenges of transcription (Jenks, 2012; Lindsay & O'Connell, 1995; MacWhinney, 2014; Oliver et al., 2006; Powers, 2005; Roberts, 1997) and is worth noting for future projects to ensure data integrity.

Conclusion

As the lead field researcher and chief transcriber of the HALL project, I was responsible for managing the workflow and the quality of transcripts. I had to balance the aims of the project against the needs of the 13 transcribers, most of whom were working on part-time contracts. I believe the transcribers benefited positively from the experience in activities after the project. One applied the skills in a new research role, another strengthened her ESL teaching skills, another used some of the data for a Ph.D., while at least three others pursued postgraduate degrees which required knowledge of transcription. All the transcribers were exposed to a deeper understanding of English language communication. They gained confidence in their ESL background and through the strengths they identified; they were able to constructively apply their knowledge to the process of transcribing.

The workshops primed the transcribers about what was expected and created an awareness about the technical and personal challenges. Regular meetings, including the discussion of conventions and the importance of performance targets, ensured on-time uniform transcriptions. Although some of the interviews were relatively short in length, the total number and their content allowed for in-depth analysis that would inform researchers and practitioners for years to come.

The future of human-curated transcripts is clouded by advances in voice recognition (VR) technology such as Google Voice, Apple's Siri, and Dragon. These technologies suggest there are alternatives, especially for people who find transcription physically demanding (Matheson, 2007); however, the use of VR technology shows that it is clumsy and labor-intensive (Tilley, 2003) and has difficulty identifying more than one speaker (Matheson, 2007); and this becomes more apparent when accents and other varieties of English are used. ESL transcribers are therefore still needed, particularly for "local" varieties of English. As varieties of English and audio interviews become more common in the research process, more attention should be paid to the resources required for transcriber training, transcribing, and quality control.

Guidelines for transcription should consider not only the conventions and rules, but also be conveyed in a tone that allows transcribers to have pride in and ownership of their work otherwise it will affect their motivation and ultimately their accuracy and productivity. As such, the keywords mined from the interviews and the key words and root words listed in Figure 5 should be incorporated into work guidelines for transcribers. A proposed work process (Appendix 2), which was used in HALL, shows the steps and tools needed to manage transcribing in a second language environment. The time spent reviewing by a second and third transcriber can be minimized by having a well-organized file management system, enforcement of formatting, and spell checking.

The role of transcription in research should be scrutinized more in research publications. Reviewers should insist ensure the reader is informed how speech was transformed into text. Researchers should explain what was included and what was discarded, how unintelligible speech was managed, what efforts were made to identify unintelligible speech, how many people checked the transcriptions, how were they trained, what was their

prior experience, and how they feel about it. By explaining the processes that answer these questions, we will be better informed on the data source and methods used to transform it. We will also feel more confident in the reliability and rigor of the data while knowing its possible shortcomings.

Project proposals need to pay more attention to the role of transcription. The time needed for transcription can be affected by the length of the interview, interlocutors, audio quality, and transcription protocols. The length can be contained by the number of questions and accurate editing of recordings by clipping the top and tail so that transcribers do not waste time on irrelevant content. The number of interlocutors and their interaction can also be restricted with attentive interviewers who can ensure all research questions are answered along with follow-ups and clarifications; they should repeat some answers or make note of unfamiliar words for quick recall during transcription.

One of the limitations of this study is my choice not to conduct follow-up or face-to-face interviews. This decision was made consciously as I had asked myself how I would feel if I had become a research subject. I also felt that there needs to be a limit to the amount of research on research. I felt that it would be unfair and almost a form of bullying to ask people who had been subordinate to me for more. The email format shielded them from a physical presence and the option to type their responses allowed them to carefully curate their responses. The open questions allowed them to provide long or short answers. If I were to do it again, I may have asked the colleagues to interview each other face-to-face and possibly conduct their own focus group. Ultimately, even with informed consent, I feel it is important to limit the amount of interaction for the sake of research. I am grateful they took part and provided the responses that have been used widely in this paper.

References

ACE. (2014). *The Asian Corpus of English*. The Education University of Hong Kong. Retrieved from https://corpus.eduhk.hk/ace/

Allwood, J., Björnberg, M., Grönqvist, L., Ahlsén, E., & Ottesjö, C. (2000). *The spoken language corpus at the department of linguistics*. Göteborg University. Retrieved from https://data.flov.gu.se/gslc/

Aronson, J. (1995). A pragmatic view of thematic analysis. *The Qualitative Report, 2*(1), 1-3. https://doi.org/10.46743/2160-3715/1995.2069

Barkhuizen, G., Benson, P., & Chik, A. (2013). *Narrative inquiry in language teaching and learning research*. Routledge.

Bird, C. M. (2005). How I stopped dreading and learned to transcription. *Qualitative Inquiry*, *11*(2), 226-248.

Breiteneder, A., Pitzl, M. L., Majewski, S., & Klimpfinger, T. (2006). VOICE recording - Methodological challenges in the compilation of a corpus of spoken ELF. *Nordic Journal of English Studies, 5*(2)*,* 161-187.

Bucholtz, M. (2000). The politics of transcription. *Journal of Pragmatics*, *32*(10), 1439-1465.

Burns, E. (2010). Developing email interview practices in qualitative research. *Sociological Research Online*, *15*(4), 24-35. https://doi.org/10.5153/sro.2232

Castillo-Montoya, M. (2016). Preparing for interview research: The interview protocol refinement framework. *The Qualitative Report, 21*(5), 811-831. https://nsuworks.nova.edu/tqr/vol21/iss5/2

Cohen, L., Manion, L., & Morrison, K. (2013). *Research methods in education.* Routledge.

Edwards, J. A. (2001). The transcription of discourse. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 321-348). Blackwell.

Edwards, J. A. (2014). Principles and contrasting systems of discourse transcription. In J. A.

Edwards & M. D. Lampert (Eds.), *Talking data: Transcription and coding in discourse research* (pp. 3-32). Taylor and Francis.

Halcomb, E. J., & Davidson, P. M. (2006). Is verbatim transcription of interview data always necessary? *Applied Nursing Research*, *19*(1), 38-42.

HALL. (2016). *English language learning histories*. Hong Kong Archive of Language Learning. Retrieved from http://www.narratives.hk

Hubbell, L. D. (2003). False starts, suspicious interviewees and nearly impossible tasks: Some reflections on the difficulty of conducting field research abroad. *The Qualitative Report, 8*(2), 195-209. https://doi.org/10.46743/2160-3715/2003.1887

Jenks, C. J. (2012). *Transcribing talk and interaction*. John Benjamins Publishing Company.

Johnson, S., & Ensslin, A. (2006). Language in the news: Investigating representations of 'Englishness' using WordSmith Tools. *Corpora*, *1*(2) 153-185.

Lampert, M. D., & Ervin-Tripp, S. M. (2014). Structured coding for the study of language and social interaction. In J. A. Edwards & M. D. Lampert (Eds.), *Talking data: Transcription and coding in discourse research* (pp. 169-206). Taylor and Francis.

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

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.

Lee, K. H. Y., & Patkin, J. G. (2016). Reading as experience: Literature, response and imagination. *Changing English: Studies in Culture and Education*, *23*(1), 67-76.

Lee, K. H. Y., & Patkin, J. G. (2017). Building the digital archive of Hong Kong English learning: Methodology, challenges and reflection. *System*, *65*, 61-68.

Lindsay, J., & O'Connell, D. C. (1995). How do transcribers deal with audio recordings of spoken discourse? *Journal of Psycholinguistic Research*, *24*(2), 101-115.

MacWhinney, B. (2014). *The CHILDES project: Tools for analyzing talk, Volume I: Transcription format and programs*. Psychology Press.

Matheson, J. L. (2007). The voice transcription technique: Use of voice recognition software to transcribe digital interview data in qualitative research. *The Qualitative Report*, *12*(4), 547-560. https://nsuworks.nova.edu/tqr/vol12/iss4/1

McLellan, E., MacQueen, K. M., & Neidig, J. L. (2003). Beyond the qualitative interview: Data preparation and transcription. *Field Methods*, *15*(1), 63-84.

Mergenthaler, E., & Stinson, C. (1992). Psychotherapy transcription standards. *Psychotherapy Research*, *2*(2), 125-142.

Mondada, L. (2013). The conversation analytic approach to data collection. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 32-56). Blackwell Publishing.

O'Connell, D., & Kowal, S. (2008). *Communicating with one another: Toward a psychology of spontaneous spoken discourse*. Springer.

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.

Patkin, J. G. (2011). *The ACE manual: Data collection and transcription for the Asian corpus of English*. RCLEAMS.

Pavlenko, A. (2007). Autobiographic narratives as data in applied linguistics. *Applied Linguistics*, *28*(2), 163-188.

Poland, B. D. (1995). Transcription quality as an aspect of rigor in qualitative research. *Qualitative Inquiry*, *1*(3), 290-310.

Powers, W. R. (2005). *Transcription techniques for the spoken word*. Rowman Altamira.

Ratislavova, K., & Ratislav, J. (2014). Asynchronous email interview as a qualitative research method in the humanities. *Human Affairs*, *24*, 452-460

Reason, J. (1990). *Human Error*. Cambridge University Press.

Riessman, C. K. (2002). Narrative analysis. In M. Huberman & M. B. Miles (Eds.), *The qualitative researcher's companion* (pp. 216-270). Sage.

Roberts, C. (1997). The politics of transcription. Transcribing talk: Issues of representation. *TESOL Quarterly, 31*(1), 167-71.

Roulston, K., DeMarrais, K., & Lewis, J. B. (2003). Learning to interview in the social sciences. *Qualitative Inquiry, 9*(4), 643-668.

Scott, M. (2008). Developing WordSmith. *International Journal of English Studies, 8*(1), 95-106.

Tilley, S. A. (2003). "Challenging" research practices: Turning a critical lens on the work of transcription. *Qualitative Inquiry, 9*(5), 750-773.

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care*, *19*(6), 349-357.

VOICE. (2013). *The Vienna-Oxford International Corpus of English* (version 2.0 online). Retrieved from http://voice.univie.ac

Walkinshaw, I., & Kirkpatrick, A. (2014). Mutual face preservation among Asian speakers of English as a Lingua Franca. *Journal of English as a Lingua Franca, 3*(2), 269-291.

Appendix 1

HALL Markups (adapted from VOICE, 2013)

1. **Pauses** were only marked if they were four seconds or longer. The length of the pause was written as a numeral and placed between two curled brackets.

S1: What did you read?

S2: (6) I can't remember

2. **Fillers** were only included with pauses or doubtful replies to illustrate difficulties in responding to questions.

S1: What did you read?

S2: (4) er (4)

S1: Can you remember?

S2: Sorry, no.

3. Interviewers were advised to avoid continuation by waiting for respondents to reply. **Continuation** occasionally occurred when interviewers tried to clarify facts.

S1: You can't remember the name of the book? Was it Harry=

S2: =Potter. Yes, Harry Potter. I read that.

4. **Context** illustrated events such as the interviewee talking to someone else to form a reply during the interview.

S2: I read this one {stops to answer phone call}

5. Each syllable of **unintelligible** speech was marked with an x instead of guessing. Transcribers could offer an interpretation for the utterance in brackets.

S2: We had to <un> xx x </un> {Sounds like "talk to"} the teacher

6. **Coinages** have been included to illustrate how learners differentiate pedagogical aspects.

S1: We had <pvc> dictationary (dictation) </pvc> ever morning.

7. **Non-English** transcription was limited to the Cantonese and Mandarin. These occasional utterances were made by interviewees who wished to clarify something in their first language before saying it in English.

S2: Er, er, er @ {Speaks Cantonese: 唔知喎 I don't know}.

8. Each moment of **laughter** was marked with @ and speech with laughter in pointy brackets.

S1: I was so embarrassed @@@ because I <@> didn't know </@> @@.

9. **Searchable media titles** were written in brackets if the interviewee uttered a variation from the original.

S1: I read Twilights {Twilight} and Hunger Game {the Hunger Games}.

Appendix 2

The Transcription Process (adapted from Patkin, 2011)

1. Back-up audio to hard drive.
2. Rename file and create list on master spreadsheet.
3. Edit audio file (if required).
4. Save audio file in common work folder.
5. Allocate transcription duties.
6. Create text file and save with same name as audio file.
7. First transcriber first draft.
8. Revisit problem areas in first draft.
9. Spell check, format check first draft.
10. Second transcriber check.
11. Return and or clarify (if required).
12. Second transcriber repeats steps of first transcriber.
13. Chief transcriber checks formatting, word count, length etc.
14. Transcription file is locked for analysis.

Possible next stages:

- Investigator questions content and finds transcription does not match audio.
- The transcription is returned to the chief transcriber for review.
- Draft versions are compared using WinMerge.
- Findings discussed and reviewed with relevant parties.

Author Note

John Patkin is a research assistant at the Education University of Hong Kong. Please direct correspondence to patkin@eduhk.hk.

Article Citation

Patkin, J. G. (2021). Transcribing in ESL: An investigation into role, accuracy, and skills. *The Qualitative Report, 26*(2), 588-609. https://doi.org/10.46743/2160-3715/2021.4438