

**How transcription is taken for granted: An analysis of transcription in
doctoral theses in education**

**Paper for the Annual Meeting of American Educational Research Association, 30 April – 4 May, Denver,
Colorado.**

Dr Christina Davidson

CQUniversity

Dr Teresa Moore

CQUniversity

How transcription is taken for granted: An analysis of transcription in doctoral theses in education

The literature on transcription over three decades asserts the taken-for-granted nature of transcription in research. Most recently, it has been claimed that transcription is neglected in doctoral training of qualitative researchers yet there are few empirical studies of transcription in postgraduate work. The article reports a pilot study of transcription in doctoral research in Australia. Specifically, the study employed content analysis to examine how transcription was addressed in twenty doctoral theses informed by phenomenography, critical discourse analysis, conversation analysis or grounded theory. Discussion considers how transcription was addressed across studies and within the particular methodologies. The study suggests the need for increased attention to transcription especially in the reporting of doctoral research.

Objectives

Analysis of transcripts is central to the work of many researchers in qualitative inquiry in education. Yet the research literature on transcription is replete with claims from qualitative researchers about the taken-for-granted nature of transcription (Bird, 2005; Duranti, 2007; Kvale, 1996; Lapadat, 2000; Lapadat & Lindsay, 1999; Tilley, 2003a, 2003b; Tilley & Powick, 2002). Specifically, it is claimed that transcription is under-discussed in journal publications (Wellard & McKenna, 2001), in methods texts about qualitative inquiry (Halcomb & Davidson, 2006), in written reports of qualitative research (Lapadat and Lindsay, 1999) and in post graduate research training (Bird, 2005; Lapadat and Lindsay, 1998).

What is taken-for-granted then about transcription in qualitative inquiry? In the literature, it is asserted that many researchers do not make their position on the nature of transcription clear, to themselves or to others (Lapadat, 2000). This omission occurs in the design of research and in its reporting. It is claimed that many researchers do not understand, or make clear, the relationship between theoretical and methodological perspectives and the development of transcripts (Coates & Thornborrow, 1999; Oliver et al., 2005). Further, overlooking this relationship has implications for analysis, findings and trustworthiness of studies (Easton, McComish & Greenburg, 2000).

Clearly, claims concerning the taken-for-granted nature of transcription in research suggest implications for qualitative inquiry in education; this is particularly so since transcripts of interviews and classroom interactions are central to numerous methodologies used in the conduct of education research. However, the transcription literature also points to the lack of empirical studies of transcription and transcription practices (Lapadat, 2000; Lapadat & Lindsay, 1998; Maclean, Myer & Estable, 2004; Wellard and McKenna, 2001). Oliver et al. (2005) assert that transcription is neglected in *certain* methodological approaches and not in others.

This paper reports a pilot study that examined how doctoral students in education reported transcription as an aspect of their postgraduate research. The study sought to address claims in the extant literature on transcription that it is taken-for-granted in reports of research. The central research question was “How do doctoral students account for transcription and transcripts in

their theses?”. The study aimed to contribute understandings of the transcription process in qualitative research. On the basis of our content analysis, we establish how doctoral students addressed transcription and we argue that transcription was taken for granted in various ways.

Perspectives or theoretical framework

Claims about the taken-for-granted nature of transcription are located as far back as the seminal work of Ochs (1979). In relation to child language studies, Ochs asserted the importance of a basic and shared notation system, the need to consider power relations inherent in representing the talk of adults and children, and the need for a close fit between the theoretical and methodological studies that inform a study and the approach to transcription that is used. Specifically, Ochs claimed that it was harder to address the goals of a study if the approach to transcription was left implicit and if it was not coherent with other aspects of research design. Although Ochs put transcription matters on the table thirty years ago, her claims stand unrefuted today (Duranti, 2007); a persistent theme in the extant literature about transcription remains that it is taken for granted (Davidson, 2009).

Transcription is an interpretive process whereby transcribers make choices about what to record (Kvale, 1996). Choices are integrally related to theoretical positions and how researchers locate themselves and others in the research process (Jaffe, 2007). Making choices or being selective is a necessity (Cook, 1990; Duranti, 1997) since it is impossible to record all aspects of recordings or interactions. A transcript that attempted to record all aspects of recordings would be unwieldy and difficult to read. An understanding related to selectivity is that transcription is a representational process (Bucholtz, 2000; Green et al., 1997). Further, transcription is understood to reflect theory and to shape it (Du Bois, 1991) as researchers “reflexively document and affirm theoretical positions” (Mischler, 1991, p. 271) during the process of transcription and analysis. Although researchers provide differing definitions of transcription (see for example, Duranti, 2007; Green, Franquiz, & Dixon, 1997; Jaffe, 2000; Mondada, 2007; Ochs, 1979), the literature shows increasing agreement on these aspects of transcription (Baker, 1997).

Recent reviews of transcription in qualitative research highlight the importance of addressing transcription in reports of qualitative studies (Lapadat & Lindsay, 1999; Davidson, 2009). Lapadat (2000) asserts that:

[T]ranscription decisions and processes employed during data collection and analysis need to be explained clearly and thoroughly in the write-up. When standardized procedures are used, a few words will suffice, but when researchers contextualize and negotiate method as a means of interpretive seeing, there is no shortcut to explicit description (Lapadat, 2000, p. 217).

According to Duranti (2007) many qualitative researchers develop their own hybrid approaches to transcription; so it would appear that transcription needs to be thoroughly documented and explained.

Method

The study employed content analysis, a method for managing and analysing data in qualitative studies. Lankshear and Knobel (2004, p. 334) suggest that qualitative content analysis is a useful tool for comparing a number of texts, either of the same type or across a period of time. Content analysis can tell us what is in the text and also highlight any “oversights” by the author/s of the text, thus allowing certain inferences to be made (Lankshear and Knobel, 2004, p. 332).

The approach was judged appropriate for our study, which encompassed a data set comprised of a number of theses. Content analysis enabled us to ask the same questions of the content of each manuscript although differing methodological and epistemological paradigms had been employed by researchers. The use of content analysis also enabled us to address an ethical dilemma about how we were to use this data and provide evidence. For confidentiality reasons we did not want to directly quote from the individual manuscripts but rather use de-identified and aggregated data. Content analysis can provide a mechanism for thematic development where sometimes the themes are inspired by a set of theoretical ideas generated and/or applied by the analyst (Gomm, 2004).

At the most basic level content analysis enables us to look at the data for recurrent instances of some kind (Wilkinson, 2004, p. 184). Kohlbacher (2006) sees classical content analysis as a quantitative method with the core and central tool being its system of categories. Likewise Berg (2004) argues, when using content analysis researchers must firstly decide on what level they will sample, that is word, item, paragraph, document and then what units of analysis will be counted. In order to know what to count explicit rules for identifying and recording of the characteristics of the content analysis is required (Berg, 2004, p. 275). While some authors view content analysis as a quantitative tool, others have used content analysis as part of an investigation of qualitative data where both latent and manifest meaning can be applied or inferred (Lankshear & Knobel, 2004, p. 333). Likewise, Wilkinson (2004, p. 187) argues that the various types of content analysis share a similar underlying epistemology whereby participants’ talk is a means to access something that lies beyond or behind this talk. Wilkinson (2004) highlights that the use of focus groups as a data collection technique has often been specifically chosen because of the group interaction in constructing talk that would otherwise not happen. It is being part of the focus group that enables specific data to emerge but rarely is this present in any transcripts produced. Transcripts that have been produced are more likely to be presented as though this interaction was one-on-one rather than as a typical group interaction (Wilkinson, 2004). O’Connell & Kowal, (1999) suggest that the use of transcription range along a continuum from explicitly stating talk as it happens in reality to searching within talk to access perceptions within that text. However it would appear that many researchers often omit or gloss over how they actually did their analysis to reach the interpretations or findings presented (Wilkinson, 2004, p. 182).

A recurring theme in content analysis is whether the analysis should be limited to the manifest content (elements physically present and countable) or extended to include latent content (Berg, 2004, p. 269). Kohlbacher (2006) highlights the added advantage of exploring both the manifest and latent content of the text. The inclusion of latent content enables a way of exploring the unspoken and expressive (Rossman & Rallis, 2003, p. 32) where texts and artefacts can provide an insight into the actions, norms and intentions of a cultural group. If we consider the doctoral researcher as belonging to a particular cultural group then by looking at what is missing within

the doctoral thesis an insight into the thinking around transcript production can be developed. For example some methodological framings would lead to the explicit use of transcripts while others tend to have a more implicit approach to the representation of data. Therefore transcription practices may reflect disciplines in the way that specific data is presented and therefore available for content analysis (Luebs, 1996).

The unit of analysis is the first decision that researchers employing content analysis need to make (Graneheim & Lundman, 2004). This unit may take various forms that can extend from a program or classroom (Mertens, 1998) to words or phrases within a transcript (Feeley & Gottlieb, 1998). Lankshear and Knobel (2004, p. 335) advocate definitional content analysis “where the unit of analysis remains at the word level, but the researcher is also interested in reading the text on either side of the focus word or phrase”. In our study the unit of analysis was “transcription”. Variants of the term included “transcript”, “transcript notation” and “transcribed”. This combination of elements (Berg, 2004, p. 274) enabled us to look for particular words such as ‘transcript’, certain items such as the inclusion of transcripts in the appendix or in the body of analysis chapters, the naming of a notation system used for transcribing and whether any issues regarding transcription were raised by the researcher.

Twenty doctoral theses were examined. These were selected from the online Australian Digital Theses Program database. Criteria for selection included: PhD degree awarded between 1998 and 2008 from a Faculty or School of Education; used transcripts of interviews or other interactions; a research design informed by conversation analysis, critical discourse analysis, phenomenography or grounded theory. As well, we tried to obtain a spread of theses across Australian universities and supervisors. Five theses were selected for each research approach (and it was challenging to find the number of theses to fit all the criteria listed above). Each thesis was assigned a code for identification (for example CA1 referred to the first of five theses that employed conversation analysis). In the end we gave each thesis a letter of the alphabet with theses grouped into methodological framing. A set of pseudo-demographic type data was collected about each doctoral thesis. This included the degree type – PhD or EdD, institution, year, theoretical perspective, methodological perspective and data collection methods. At the beginning stage both of us looked at each thesis in turn. This was so we could both get a feel for the data and develop a shared meaning around the questions and therefore methodological patterns present within the thesis document.

The following broad procedure was then employed as part of the initial coding and categorising of the data:

Step 1: read through each thesis marking relevant content initially by looking at the table of contents, methodology chapters and appendices.

Step 2: when looking at “transcription” a decision was also made about what else might require coding from the accompanying phrases on either side of the unit of analysis

Step 3: a content summary sheet was developed and used to record information from each theses

Step 4: revisited the chunks of data on summary sheets as definitions were refined and reiterated

Step 5: interpreted results across all theses and within the four methodologies.

In the development and reporting of our study we have negotiated and addressed common limitations often cited with the use of content analysis. These include: over-extended inferences, possibly disregarding the purpose of the text or over-reduction of data.

- Over-extended inferences: we would agree that it is possible to “read” too much into some inferences coming from the data however in our case we are looking for specific instances of how various PhD candidates have used, written about, referred to transcription as a practice in the thesis product. Therefore we would content that the candidate has either done this or has not done this at the basis level of data collection
- Possibly disregarding the purpose of the text such as why it was produced, by whom and for what audience: this point raised issues for our analysis in that we needed to consider the discipline of the thesis product. In the case of a thesis using Conversation Analysis, the candidate would have necessarily included a transcript as this is fundamental to the analytical method whereas this is not necessarily the case in the various forms of analysis employed within other methodologies
- Over simplifying the data or over reduction: we would consider that if the candidate did not include a transcript or discuss transcription then there was no further reason to investigate that particular thesis product however if the word transcription appeared then we were looked at how this was manifest in the thesis product.

The purpose of the text often assists in alerting the reader to decide in how to “read” the text (Bauer, 2000 in Lankshear & Knobel 2004, p. 334). In the section that follows we present findings that are exploratory in nature.

Findings

When we started this study, our reading of the transcription literature suggested that we might find clear lines of division when examining how transcription was approached in studies informed by grounded theory, phenomenography, critical discourse analysis or conversation analysis. The findings showed that approaches to transcription and its reporting in PhD theses are not as clear cut as some of the literature suggests in relation to particular methodologies. That is, there were individual differences in the reporting of studies informed by the same research approach. In order to preserve the richness of these differences, we present our findings as (1) common features within a methodology, and (2) features where variation occurred. We also found that some aspects of transcription and its reporting were taken-for-granted across all methodologies and these are also outlined here. In all cases, findings will encompass some of the following categories that we delineated: transcription and its relationship to theory and methodology; transcription and data; transcription and analysis; doing transcription; transcription notation; use of transcription and transcripts in the thesis; use of transcripts in the research process; addressing transcription as an aspect of trustworthiness.

Grounded theory

Reference to the transcription process emphasised getting down what was said accurately. So we find use of words such as “exactly” (thesis D), “in full” (thesis A), “accurately and in full” (thesis D), “verbatim” (thesis E), “accurate record” (thesis E), and “accurate transcription” (thesis E). Researcher D provided a list of what was transcribed that included “false starts, repetition, redundancy, recasts, phatic language and occasional errors of a non-systematic kind”

(thesis D, pp. 90-91) although this was the only specific detail provided in the theses that addressed more than words. Researcher B wrote that “each tape was played, and replayed many times”. Other researchers emphasised the need for expediency in transcription. The words used in relation to transcription place emphasis on recording words as they were said. Since no researcher defined transcription in other ways, we take it that these researchers were drawing on views and understandings of transcription as recording words that were said in interviews or classroom interaction.

According to the researchers, transcripts were developed from recordings and became “the data for Grounded Theory coding” (thesis A, p. 78). Four of the researchers referred specifically to transcripts as data to be analysed. This confirmed for us, understanding of transcripts standing for recordings and being data, although for one researcher, it was important to return also to the recordings during the analysis since:

The work of Kvale (1996) also advocated caution in the use of transcripts in analysis for “interviews are living conversations”. In order to maintain the “living” aspect of the interview, the researcher used video collection of data, and these were constantly reviewed during the period of analysis (thesis C, p. 91)

In describing the role of transcripts in analysis, all researchers provided details about how the analysis of transcripts was undertaken. In at least one case, a researcher makes clear that transcripts were analysed multiple times. Analysis involved reading and re-reading transcripts and writing on the transcripts. In this aspect –analysing transcripts –researchers provided a deal of information, thus appearing to emphasise the analysis of transcripts rather than the transcribing of them.

Two researchers identified the relationship between data collection and analysis as dependent on transcription for the generation of further questions and further data collection. This raised the issue of time needed for transcription and how it impacted on the feasibility of further data collection driven by transcripts of interviews or questions that have arisen from them. Clearly, getting to transcripts quickly is central to applying the methodology as it has been described and envisaged.

Another researcher (thesis A) used transcripts of interviews to provide for further comment from interviewees. So, following the production of the transcript, it was sent to the interviewee who was able to elaborate on points ie. to add in additional information in a blank column that was included along the side of the transcript. This information was then added to the transcript excerpts used within the theses, although indicated with italics. Researcher D provided the explanation that transcripts were edited lightly for the thesis. These practices provide food for thought about what *is* the transcript, what *is* data and what *is* presented in the thesis.

The researchers also varied in relation to their inclusion of excerpts and examples or even entire transcripts and how they referred to these. In one case (thesis B), one hundred pages of transcripts were included as an appendix in the thesis. Other researchers made reference to “quotes” or excerpts or “quotations” from the transcripts. Lines or turns were not numbered in any of the transcripts, therefore appeared not necessary for the effective use of the transcript materials.

Apart from presenting how grounded theory researchers addressed transcription and transcripts it is also important to identify what wasn't addressed. Researchers working within grounded theory did not present a theoretical perspective on transcription or transcripts as partial, selective, representational or interpretive, for example, apart from researcher C's comment on keeping interview data "living". There was no mention of how theoretical perspectives on transcription could be brought to bear on the way that transcription was addressed and conducted in studies. Researchers did not use a notation system, or rather did not name a system if they had employed one. Further, researchers did not address transcription as an aspect of the trustworthiness of their studies. No researchers raised difficulties or issues to do with actual transcription (in the sense of getting talk and interaction down onto the page). There was very little reference to literature related to grounded theory and transcription although one researcher cited literature which confirmed the method of analysing transcripts as data:

Strauss and Corbin (1990) outline progressive stages of analyzing data as (1) open coding in which all the data are broken down (by means of writing associated words all over the interview transcripts); etc" (thesis D, p. 19)

There was one researcher (thesis B) for whom transcription was seen as integral; this was because the change from sound and image into print / text is an integral step in moving raw data into conceptual thought. This step is seen as a way of making the relationship between data collection and initial data analysis. Having a transcript, it can therefore be argued, enables the researcher to have a 'useful' and comparable object – ie the transcript contents where coding and thematic development takes place through a constant comparative method or technique.

It was noted during our analysis of thesis content (thesis B) that some content within individual transcripts could reflect the influence / interest of academics and supervisors within the candidate's learning and scholarship community. In this particular thesis there was evidence of conversation analysis notation however it lacked any discussion pertaining to its inclusion.

Transcripts could be perceived as a common denominator or a way of re-constructing field notes, videotapes, audio-tapes and documents into one format – that of written text or into the same format. Italics are a quoting device and a visual cue to indicate that the contents are from a data source rather than words of the researcher / writer. In doing so the content appears to take on the appearance of written wording rather than talk within quotes, thus indicating perhaps this comes from the changed data format, ie the transcript not the original raw data.

Phenomenography

The analysis revealed that the phenomenographic researchers provided a deal of information about their use of transcripts; much less information was provided about the transcription process itself. Phenomenographers' descriptions of their analytic process gave emphasis to working with transcripts as 'doing analysis. Together, researchers' comments portray a lengthy analytic process that involved a lot of reading and re-reading of transcripts (thesis F, thesis G, thesis H, thesis I, and thesis J). One researcher also included watching recordings in tandem with reading of transcripts (thesis G). The analytic intent was a focus 'on the meaning embedded in the utterances instead of the specific words used by the individual participants' (thesis F, pp. 98-99) and "going beyond particular words and statements" (thesis G, p. 112); with this being done by "looking into transcripts" (thesis G).

The word ‘transcript’ appeared numerous times in sections or paragraphs in two theses (thesis G and thesis J) where researchers discussed an analytic issue related to coding and developing of conceptions and their descriptions that is central to phenomenography. The issue relates to how researchers work with transcripts during analysis: should sections of transcripts be extracted from their original place in an entire transcript of an interview or remain within the transcript? The matter of whether to take the transcripts out of the context of their use, or not, gives rise (in particular) to procedural information about the analytic process which relates to transcripts and their analysis.

The word ‘transcription’ appeared less frequently but occurred in tandem with “verbatim” at some stage in every thesis (thesis F, thesis G, thesis H, thesis I, and thesis J) when the transcription process was described. The co-occurrence of the words suggested the importance of “getting down words” rather than recording the ways that words were said or other physical actions that occurred during talk. Researcher J, however, provided elaboration on the process of transcription in this way:

Participants’ words were transcribed verbatim for meaning. Other sounds, like pauses or noticeable changes in tone of voice, that contributed to meaning were also included in the transcript. Words used to describe these sounds were placed in brackets so they would not become confused with the actual text (thesis J, p.x)

This researcher also described how transcripts were edited for the thesis and provided examples in her thesis to illustrate. This way of addressing transcription illustrated what transcription meant for the researcher and also provided description of a practice that related to the presentation of data and its analysis in a thesis.

Original – Christine: Um, well, I’ve got to sort of think about my three classes this year if you’re talking about me personally and at various times, some kids will be more engaged than other kids, but I can’t think of anybody who is always not engaged....”

Edited version-Christine: ...At various times, some kids will be more engaged than other kids, but I can’t think of anybody who is always not engaged....” (p. 143)

This elaboration was unique in the phenomenographic studies, although was not the only study that provided a deal of information about the presentation of transcripts in theses.

Two other phenomenographic researchers (thesis F and thesis I) provided details about their use of translation in transcription. Both researchers conducted interviews in Chinese, did the original transcription in English and then at a later stage, developed a transcript in English. Both researchers gave consideration to the stage when an English translation was necessary: both did the analysis on the Chinese transcripts and developed transcripts in English for the thesis. This was to allow for ‘getting at the meaning’ in the original language of the interviews. One used only excerpts from the English transcript in the thesis; the other provided both Chinese and

English in the excerpts used in the thesis. Both researchers presented this as a translation issue rather than a transcription issue.

Despite differing positions on transcripts and the analysis of them, one thing is clear: transcripts are data. Clearly, within this approach, researchers are “looking into transcripts to discover the particular ways in which people understand the phenomenon” (thesis G, p. 20). It is not the recordings that are data but the interviews that “are fully transcribed and then the data is analysed in several phases” (thesis H, p. 123).

In the first stage, the researcher analysed the transcriptions and identified sets of categories of descriptions meant to describe the key aspects of the variation within the set of transcripts as a whole. The analytical process was iterative. The researcher then reviewed the transcripts according to the designated categories, revised categories and the relationships, and revisited the transcripts and categories until stable sets of categories and relationships were developed” (pp. 128-129)

As the interviews had generated hundreds of pages of data, the computer software program NVivo was utilized, primarily as a data management system” (thesis J, p. 144).

It is important to tease out here that, although some researchers directly name transcripts as data, others infer that transcripts are data as they write about their analytic process.

Overall, the phenomenographic researchers did not preserve the original features of spoken language in their transcripts. All researchers used quotes in their theses that removed distinctive features of spoken language so that transcript excerpts resembled written text rather than talk as it occurred in the interview. As well, although not named specifically in relation to translation, clearly producing a thesis in English presumed an English speaking reader, while providing dual transcripts preserved the original words for a Chinese speaker (if only the researcher). No researcher made reference to the use of a specific notation system although transcripts shared features in common: they used punctuation from written language, for example. Gestures and other non-verbal actions were not recorded. These aspects of transcripts produced written-like texts. Researchers did not write about transcription from a theoretical perspective nor draw on literature related to transcription in qualitative research. Two of the researchers (thesis F and thesis G) employed member checking and wrote about it in relation to accuracy and getting words down on the page. This, perhaps, affirms the importance of words to a phenomenographer rather than how those words are expressed.

Critical Discourse Analysis

As we moved into the critical discourse analysis realm there was the beginning of an acknowledgement of a theoretical positioning by stating that transcripts are constructed and therefore more than translation (thesis K). Transcripts were described as particular kinds of objects and re-representation of data:

“The data within this study were re(presented) by such diverse texts as printed transcriptions, audio recordings, scanned images, photographs and written texts” (thesis N, p. 113).

Issues of power were highlighted with this relationship being on different levels. More specifically, the way that transcripts and transcription is talked about is consistent with poststructuralist theory that is being drawn upon. Information has been provided about how the transcription was done and about the way speech was recorded. In particular, it has been made clear where excerpts or “snippets” came from in the data in a systematic way. Within the CDA methodology / approach there has been emphasis made of the importance of being systematic in the approach to transcription across all the spoken data.

The transcripts have been described to the respondents as ‘word processed’ transcripts when they are returned for member checking. Little discussion arose however around power in relation to changing transcripts after the respondents had read and commented on what they regarded as an interview and spoken data. In this sense, on the one hand, we have a manipulated text (author) and on the other, a copy of what has been perceived as an interview minus the features of an audiotaped interaction (respondents). There was no reference to discussion regarding the act of transcription with the respondents who are the original source of the data. It could then be asked – how does this action influence what the respondents’ state after reading the transcript, how might this change the meaning and later interpretation? This is graphically highlighted where one respondent who was described as a teacher states “I was disgusted with myself. I read the transcript and thought I sounded like a Year 3 student” (thesis M).

For the CDA theorist in theses examined, the transcript “becomes” a text to enable textual analysis, however it is not the document as a whole that is being analysed, merely the content of the document. What is meant here is that the author is doing the analysis on the content of a text rather than looking at the content, layout and inclusion of images as in a more traditional document textual analysis. Although one author named the act of presenting transcripts as necessarily biased: “The transcripts were considered to be a necessarily biased re(presentation) of the data” (p. 101) this appeared to refer to layout. Then this layout may have been critiqued as part of the ‘power’ lying within the format and content of the text to be analysed. The researcher talks about the resultant transcript as a more effective reference tool than the original source.

“The transcripts were seen as an effective reference tool, rather than an original data source” (p. 101).

Here the transcript is clearly a re-representation of the data as the source of the analysis; although the act of transcription is not regarded as analysis. There has been some interplay between the original and the re-representation, so additional notations – meaning here, notes being added to the piece of paper or electronic page, more than analysis being developed - could be made by the researcher. Similar to memoing in GT methodology, these notations are becoming part of the actual analysis process. In the beginning, the transcript appears as a flat representation of the original spoken data that is then built upon as the author develops further in-sight into the analysis of what was previously spoken data. It would appear that while the written format is read, the author listens to the original recorded data and makes notes on the transcript document. The author explains their own positioning on the level of analysis of the transcript thus informing the reader of why the resultant analysis did not include specific indicators of CA-like notations. There is an explanation of the source of quotes – the author explicitly notes the precise location of segments within the transcripts therefore we would argue that transcription and transcripts are

addressed much more evidently in this thesis (thesis N) than in the other CDA examples examined.

Although the description of transcripts is handled in ways in keeping with the methodological approach where transcripts are described as an artefact, that data has been transformed with this 'act' of transformation has now been both acknowledged and perceived as the basis for analysis. It is pre-preparation for analysis. It is the constructed transcript that is then analysed, despite the claim within the thesis by researcher K where Gee (2005, p. 88) is used to substantiate "that constructing transcripts is already an act of analysis". We would argue that in this case there is an ambiguous interpretation of this quote to substantiate something different. Researcher K does acknowledge editing and does explain the interaction and manipulation within the wording of the transcripts. The taken-for-granted aspect here lies in the use of a notation system where the usage of some symbols remains unexplained. There is a list of excerpts indicating quotes or sections of the transcripts that have been used within the thesis document. Within these excerpts there is a closer resemblance to spoken data with some of the features of spoken data preserved. There has been some argument made to explain the slippage of dialogue, data, transcripts and conversations as though the value and notion of each are the same; it could be argued here that Researcher K has attempted to re-represent data in a form that imitates spoken data in the raw form, without the benefit of understanding why. For this researcher, similar to the GT theorists, the original data is transformed and forgotten while intense interest centres on the constructed but hybridized transcript without much discussion of the act of transcription.

There was a highlighting of power in being able to manipulate data though an act of transcription acknowledging some implicit knowledge that accompanies transformation of data and text (thesis L).

"One final point is that the power I carry as researcher extends beyond the interview and transcription protocols. I also had editorial discretion in terms of deciding what to include in my thesis chapters and choosing how to render the accounts of their experiences."
(thesis L, p. 68)

While Researcher L has stated that transcripts were edited for fluency, thus demonstrating a particular kind of cleansing, some resemblance to spoken talk remains. Data sources employed within the thesis are indicated through a 'universal' numerical system of participant ID, interview number and line number. More than one person did the transcription; demonstrating the act of typing was about direct translation of spoken word to written word without any thought given to the translation of underlying inferences that occur within the act of speaking. Much care was given to anonymising the respondent, however less care was shown to the handling of the words and talk actually done by the respondent, as illustrated by the reference to editing. There appears some tension between the notion of verbatim and the respondents' view of their own speech without the Researcher detailing differences between spoken and written text with the respondents. It could be asked again – why were so-called verbatim transcripts returned for member checking then edited? The author begins the analysis process by re-reading the transcripts but does not indicate the listening of the original audiotaped data.

In the theses examined, the act of transcription transforms spoken data into a transcript, which then re-appears as a text. An example of this is thesis M, which supports our arguments of transcripts magically turning into texts with the researcher explicitly referring to “written transcripts of spoken texts (interviews)”. A good description of how the researcher has used segments of individual transcripts within the thesis has been provided. The act of transcription here is used as a tool, similar to GT theorists we examined earlier, of bringing all data artefacts into the one form:

“Fairclough’s version of critical discourse analysis also enabled the examination of all ‘texts’ within one conceptual framework. As will be discussed in more depth in Chapter 5, data collection for this study involved a range of texts, including written transcripts of spoken texts (interviews), field notes, newspaper articles and school documents, all of which involved language use and were thus the products of social interactions of various types. All could be considered as texts, as examples of discursive practice and as instances of social practice, as conceptualized by Fairclough’s model (see Figure 1)” (thesis M, p. 28).

There is however reference to the non-neutral process of transcription, in keeping with the methodological framing where there is sensitivity to power relations, not only between the researcher and respondent, but also between the researcher and the ‘texts’ that are being interrogated. Power, therefore lies with the wording; how words are used in sense making and the objective of the text. In this case, researcher M does indicate the ‘messy’ aspect of spoken talk; while not every ‘um’ is illustrated the quotes do retain features of spoken data. However, this practice was not evident across other theses.

Conversation analysis

Conversation analysts all employed a notation system (which four of the researchers named as Jefferson notation system). The system has been central to the development of CA and so its appearance in the theses is not surprising. The use of the word ‘notation’, ‘system’ or ‘convention’, together with transcription, showed that the use of a system is integral to transcription in the approach (such that it could not be transcription for CA if a notation system wasn’t used). All researchers included a key for their system as an appendix to the thesis, therefore they provided the reader with access to information about the system to enable their reading of it. As well, every researcher made use of the literature from conversation analysis where the transcription system has been outlined previously (for example, Atkinson & Heritage, 1984; Psathas, 1995; ten Have, 1999). In a sense then, the matter of transcription was already “on the table” for neophyte researchers becoming familiar with the approach through reading seminal literature on the development and methods of CA.

Four of the researchers (thesis P, thesis R, thesis S and thesis T) also discussed aspects of the notation system in the body of the thesis and delineated and discussed modifications that they had made to the notation system during transcription and the development of the thesis. Modifications to the system are not unusual in CA –usually to encompass aspects of a study that might not have require a level of detail that the Jefferson system does not provide (see for example, Hepburn’s development of symbols to represent different types of crying). Two of the researchers, in our study, discussed and explained the modifications. For example:

In transcription, I represented the talk of the participants using modified orthography. As Ochs (1979) notes, this approach roughly captures the way talk sounds. I used phonetic representations when I transcribed children's playful sounds or unintelligible speech (Ochs, 1979). The following extract illustrates some of these conventions (thesis T)

The third step in the transcription of the episode of the girls' interaction in the sandpit involved the development and use of symbols. I developed symbols that were suitable for actions that were common in the sandpit, such as digging in the sand and sieving the sand ... These symbols were placed in an ordered manner (symbol, space, symbol, space, etc.). However, the actually digging and sieving in which the children engaged was not so methodical... (thesis T, pp. 71-72)

Here, modification of the system is specifically related to the data to be collected and the representation of it in the transcript; this has necessitated modifications to the Jefferson system.

From the perspective of ethnomethodology, data are the recordings and transcription of recordings an integral aspect of analysis. This is an important understanding since it requires, therefore, that analysts should do their own transcriptions. Since these transcripts are very detailed, producing them is very time consuming, requiring numerous listening/reviewing of recordings. Where we found differences in our corpus for ethnomethodology, was in the ways that transcripts and transcription were referred to in relation to data and to analysis. So, for example, some researchers mainly referred to the transcripts as data where others referred to transcripts and recordings as data. Word usage occasionally introduced finer distinctions related to what was considered to be data, for example, one researcher referred to recordings as 'raw data'. For some, analysis was something done to transcripts whereas for others transcription was analysis. For example:

the data for this project include only audio-tapes and transcripts of actual consultations recorded *in situ* (thesis P, p. 165)

So, in places it was possible to find data written about as both recordings *and* transcripts of those recordings. For others, data became the transcripts and recordings mentioned less frequently and not in tandem with transcripts as data. Transcripts or 'transcribed recordings' were sometimes referred to as accounts, where accounts encompassed phenomenon *in situ*. For example researcher Q refers to "accounts of 'what counts as reading' in particular settings" (p. X). Another researcher (thesis P) refers to transcripts as 'interpretive accounts' that need to be used together with recordings during analysis. The matter of CA's take on what is data has been the focus for some discussion (see for example, Mischler, 1991).

Another shared aspect in the PhD theses of ethnomethodologists was detailed explanation about what was recorded and transcribed. Since application of Jefferson notation is very time consuming in the transcription process, not all recordings are necessarily transcribed. So in their theses we find 'transcript' and 'transcription' being addressed in relation to what was recorded and what was transcribed. These references are both procedural (explaining what was recorded) *and* a form of accountability whereby researchers made clear how much of the data was analysed

and whether the analysis was single case or materials taken from across the corpus of recordings (or transcripts, for some). Both methods are used on conversation analysis.

In all cases, transcripts were appended to the thesis. These provided examples of full transcripts of individual recordings from which episodes had been selected for recording (thesis P, thesis Q, thesis S and thesis T) or were all the transcripts produced from which examples had been taken and analysed in the thesis (thesis R). In the case of thesis R, the inclusion of a large amount of transcribed materials required a second volume for the thesis.

For conversation analysts, transcription (and transcripts) was integral to the trustworthiness of their studies. This was strongly addressed across the corpus, although differently, and included the following points or views: audio recordings always referred to during analysis –so not just analysis of transcripts (thesis P), recordings and transcripts offer a reliable record of events (thesis Q), the public nature of transcription and transcripts (thesis R, thesis S and thesis T), quality of transcripts (thesis S), use of conventions (thesis S), agreement on transcripts through sharing recordings and transcripts of them with others (thesis S), acknowledging transcription as a selective process (thesis R), and adequate transcription of audio data (thesis P, thesis S and thesis T). The latter point, made by three researchers, illustrates the importance of the CA literature in ethnomethodological work. The researchers draw on and cite the work of Perakyla (1997):

the key aspects of reliability involve selection of what is recorded, the technical quality of recordings and the adequacy of transcripts (p. 206)

The use of this work, in theses sections addressing trustworthiness, illustrates the way that the literature on transcription can be used or drawn on in a thesis to good effect.

Across theses

In our analysis of theses we noticed that “transcript” or “transcripts” was used in two specific ways. Use of transcripts across the set of all theses encompassed use in the research process and the specific subset of the production of the research thesis. Uses included: producing data from a recording of interview of classroom lesson; as a member checking device; as a device to elicit data or further data. For example, in one CA theses use of transcripts was an important data collection tool in this intervention study. Therefore it is the word ‘transcripts’ rather than ‘transcription’ which appears a lot throughout the thesis. This is accompanied by reference to research participants’ reading transcripts and using them to alter their teaching.

Within theses, transcripts were: evidence of analysis; illustrative of categories, reference point for doing analysis within a chapter; a check provided for the reader. As such the use of transcripts or transcription was often directed at helping the reader to know how to read the thesis. For example, by pointing out that a particular transcript “showed something” or “was evidence of something”. While not the focus for our study –how transcription is dealt to in theses-the use of transcripts in theses would be an interesting topic for further research. We note Luebs’s (1996) study of use of transcripts and excerpts in journal articles as an interesting source for considering use of transcripts in writing generally or specifically in documents that report research studies.

Analysis of the twenty theses provided little information about how researchers actually *did* transcription. For example, doing transcription often requires or makes use of various technologies –from machines to computer programs. Yet researchers did not write about or mention use of technology. Nor did researchers outline steps or stages. For example, did the researcher make several passes through the recordings, gradually refining the transcript? Did the researcher transcribe a small section and then move on the next when satisfied? Even in the CA theses we found no information about how the notation system was systematically applied during transcription.

The literature on transcription is remarkably taken-for-granted in doctoral research across all theses. Although researchers methodically showed their understandings of the research process in relation to interview conduct and procedures by citing relevant literature, they did not give this same attention to transcription. Only two researchers cited the seminal work by Ochs, and none cited recent transcription literature from journals that address qualitative inquiry and methods (see for example, Bird, 2005; Grundy, Pollon, & McGinn, 2003; Lapadat, & Lindsay, 1999; Poland, 1995; Tilley, 2003a, Tilley, 2003b).

Further, few problems or issues about transcription were raised in the corpus of twenty theses as they were *experienced* by individual researchers. Since, transcription is widely acknowledged as a time-consuming and difficult process –frequently hived off to paid transcribers by experienced researchers-the absence of information about this aspect of the research process appears to be a *noticeable* absence (Sacks, 1995). In the CA work we found explanation for the development of new notation symbols to address aspects of specific studies, and these did address issues pertaining to recordings aspects of data. However, we did not find transcription addressed as a process that encompasses on-going issues for researchers working from this perspective.

Discussion

The experience of reading, re-reading and analysing twenty theses, was a useful and interesting process. It was driven by a “motivated interest”; that is, we were interested to see how researchers documented their use of transcription in the research process given comments in the transcription literature about the taken-for-grantedness of transcription in qualitative research. Our own work as qualitative researchers within education, and as academics working in the area of postgraduate studies led us to investigate transcription in the reporting of education research.

In the discussion that follows, aspects of transcription are considered in relation to (1) different practices that researchers engage in, in the use of transcription and transcripts, and reporting of their use (2) slippage when recordings become transcripts (3) awareness of literacized practices (4) the seamless process of transcription (5) Reporting transcription as an aspect of study design and research logic.

Researchers are doing different things with transcription and with transcripts

The analysis revealed that researchers ‘doing transcription’ are doing different things. This is evident in many aspects of the research process and its reporting. The difference is not just about how transcription is used; it encompasses the ways that the transcription process is conceptualized, described, reported and justified. While, we have been determined to show the

richness of the variety of practices, rather than the ways that transcription might deviate from some ideal version of what transcription is, it is nevertheless useful to consider how an aspect of the research process can differ in ways that are related to theoretical perspectives of studies, what is considered to be data and analysis, and how transcription should be conducted and transcripts used, in the process including its reporting in the thesis document. We suggest that consideration of these differences provides a compelling argument for qualitative researchers to reveal more of their taken-for-granted practices in around transcription.

A comparison between GT and Phenomenography illustrates our point here: In the GT examples recordings were quickly transformed into a transcript. This was routinely done as part of the initial analytical process. In this case the transcript serves as an interim step between getting a volume of words or utterances from an initial series of phrases / paragraphs into a series of developmental codes that in turn become categories. The transcript quickly ‘disappears’ once open coding is completed and categorization begins. This proceeds until the point of saturation – saturation being indicated when no new categories emerge as further interviews are completed. In essence the original data form disappears and new data forms appear. This is seen as unproblematic and as though the new ‘transcript’ is the same as the original raw data. Here transcription is a means to an end and again the action of transcription is seen as unproblematic with an underlying assumption that nothing is gained or lost through the process of change.

Phenomenographic researchers on the other hand are very interested in transcripts, although less interested in transcription. They describe their analytic process in great detail, giving emphasis to how they work with transcripts. The on-going issue in phenomenography about whether to take the transcripts out of the context of their use, or not, gives rise (in particular) to much discussion about transcripts. Far less attention is given to transcription. Phenomenographic researchers, while focusing on transcripts, also write about analysis as ‘going beyond words’. Unlike with GT data however, with the movement from interview to transcript within phenomenography there is the retention of the original data, albeit in written form, which forms part of the analytic process. While the transcripts are central to the analysis, the development of these transcripts received little attention. What we can say here is that phenomenographic researchers are very interested in transcripts whereas GT researchers are very interested in their comments recorded on transcripts. The latter seem to leave the transcripts behind rather quickly. Phenomenographic researchers talk about transcripts a lot because they are the data. One conclusion that this brings us to form is that clearly, transcripts and transcription do not mean the same thing to researchers working within particular methodologies within qualitative inquiry.

It is also possible, to find differences in approaches to use of transcription within research methodologies or approaches. The example of CA is a useful one, primarily because CA researchers have available a transcript notation system. Not all the researchers acknowledged using the system; those who did still developed transcripts that looked different to each other. That is, it was possible to infer that researchers had applied the notation system differently. What resulted was a kind of “fingerprint” whereby transcriptions bore the mark of the specific researcher who had made them, despite drawing on the same notation system and in a systematic way. We attribute this to differences in how the transcription process was approached in relation to the application of the system and the development of a transcript. CA researchers did not give detailed information about how they went about developing their transcripts apart from use of the

notation system. What we do know about CA researchers however is that they develop their own transcripts – we just don't know how they systematically approach it (individually). There is a deal of commentary on transcription in the CA literature, however it does not necessarily describe how to go about the transcription process systematically. We have found in the literature that ten Have (2007) provides a guide to how a researcher from this perspective might approach doing transcription:

My general suggestion for making transcripts is to do it in 'rounds': start by putting down *what* has probably been said in standard orthography, and add the various details concerning the *how's* later, one type after the other. One can, of course, make a note of remarkable details in earlier rounds than those in which one concentrates on a certain type, but it proves a good practice to focus on particular kinds of phenomena one after the other, for instance 'intonation', 'pauses', etc. (2007, p. 111)

Note that ten Have provides a description of how to go about incorporating symbols in a manageable fashion. Of course, this is less useful if a researcher is not specifically applying a notation system.

Slippage: Interview/Interaction to Transcription to Transcript Interview = transcript

Our reading of certain sections prompted two questions related to transcripts and transcription: What is the relationship between transcription and data and what is the relationship between transcription and analysis? Here we consider 'slippage' in relation to data and to analysis that we observed.

It would appear that with GT researchers there is slippage that occurs where interviews/interaction become transcripts without the need for explaining this transformation; meaning here that the content of interviews and/or interactions are almost automatically referred to as transcripts regardless of whether there has been minimal or in-depth transcription acts to the original data form. In this form the transcripts rather than the original audio or video recording are then believed to be the data. Sound and sight is transformed into text and while the spoken word and utterances/action are markedly different to written text, the resultant transcript has the appearance of written text, and once transcription is complete the raw data is put aside or disappears. Researchers in the studies examined have not explained their approach to transcription in great detail, thus implying a taken-for-granted aspect to this section of the research process. Therefore it could be argued that transcripts are made with the assumption that spoken and interaction data is the same as written data. Further, the words 'transcript' and 'interview' may be used interchangeably and the original form or the raw data is left behind or forgotten. This transformation is not referred to again, other than to indicate a source of useful quotes that in turn illustrate analytic points within the thesis.

While the act of transcription is left unexplained this slippage and movement between forms of data will remain both problematic and taken-for-granted. This slippage of terminology and of action where transformation of data is smoothed over does not make explicit or demonstrate how the researcher arrived at this point. This is graphically illustrated in the ways that phenomenographers referred to some aspects of handling data. A common theme emerging from our examination was the act of transcribing interviews "verbatim". The transcripts would then be

returned to the respondents for checking. Since, these researchers had transcribed verbatim, there this practice appeared to be a superfluous one. What was there to check? The use of member checking within theses often appeared to generate more data which were incorporated into the transcript as if they had always been there.

Slippage in the discourse analysis methodology theses revolves around the usage of ‘text’ and ‘transcript’. Likewise the use of the word text and transcript appeared interchangeably within the CDA theses examined in much the same way as interview and transcript became interchangeable within GT theses examined.

The seminal work in CA strongly emphasizes the importance of recordings as data and that recordings remain the data. We found this position echoed in the work of some of the researchers. At the same time, procedural descriptions within the thesis evidenced some slippage whereby transcripts were referred to as data. We also found that while some researchers emphasized the importance of analyzing the recordings and transcripts together, others “left the recordings behind” in their analytic chapters and reference to transcripts.

Awareness of Literacized transcription practices

In the findings section, we noted the absence of naming of a notation system in theses from grounded theory, phenomenography, and critical discourse analysis. While we are not suggesting that researchers within these approaches should employ the one system (as conversation analysts do, for example) we were interested in how transcription ‘got done’ given the absence of a particular way of noting features of talk and interaction in interviews and classroom interaction. That is, what aspects of spoken language were preserved through transcription and what features of written language were introduced. From researchers’ references to features of talk that they recorded, and omissions, it was apparent that researchers were drawing on literacized (Bucholtz, 2000) practices during the transcription process in place of a specific notation system. Bucholtz explains literacized transcripts as those where features of written talk are emphasised rather than preserving features of spoken language. She also refers to these as naturalized transcripts and transcripts that preserve the features of spoken language as denaturalized and points out that readers of transcripts unused to seeing spoken language written down in this way will find the transcripts odd looking. This, in fact, was noted by two researchers during use of transcripts for member checking. As Baker (1997, p. 114) notes “the auditory has been imbued with the character of writing”. Our comment in relation to literacized practices is that researchers appear to take-for-granted that they are in fact imposing a system onto recordings of spoken/visual data that in fact alter it rather substantially, and frequently do not acknowledge this. Even the use of ‘verbatim’ as a way of getting down talk is subject to literacized practices such that talk is altered through the use of punctuation (for example) – one is getting more than words down on paper.

With this in mind it could be argued further, that researchers draw on their understandings of literacy conventions, which are left unexplained and thus taken-for-granted. This best illustration of this kind of practice is the use of ellipsis dots (...) to represent something left out, or the application of other symbols which have a specific meaning in relation to English print, but not necessary in relation to transcription. We did not find, for example, researchers stating that they had applied the principles of English punctuation and written language in order to develop their transcription. In the CDA these for example where notation has been used an assumption has

been made of a shared understanding, or universal claim within the local community of scholars, for example, readers of the thesis, of the use of this notation. No other discussion or explanation has been provided about the interpretation of the use of notation. Therefore it could be argued that this is a further example of slippage a literalized practice.

The seamless process of transcription?

The production of a thesis frequently results in, or relies upon, tidying up of the research process such that its description is more seamless than the experience actually was. We suggest that this is particularly the case with transcription because it is presented, reported, frequently without consideration of decisions made and lessons learnt –that is, issues that confronted the researcher during transcription are omitted from the thesis such that the transcription process is represented as straight forward (though perhaps time consuming). We saw this across theses in this pilot study – discussion of issues that confronted the researcher was rare. Transcription was presented as straightforward, unproblematic and unremarkable. This is interesting given the literature which extrapolates many of the complexities of producing transcripts. Use of the word “verbatim” was one method for presenting transcription as straightforward and accurate. Reference to “member checking” was also used to imply attention to correctness or accuracy in transcript development. Even in conversation analysis, where transcripts received a lot of attention, transcription was not discussed as a complex practice involving selectivity and issues to do with representation.

In order to illustrate how the challenges of transcription might be addressed we cite a researcher from conversation analysis (thesis S) who provided a detailed examination of the challenges of transcribing video recordings in her study:

The third and fourth steps arose in response to my increasing awareness of the overlapping layers of the interaction. A problem arose when I realized that the transcript did not capture or describe the participants’ actions accurately: much of the action was not being depicted, as often the children were digging or sieving sand as they talked. Using the Jefferson method of transcribing, action was described in double brackets after the spoken words, and hence did not show when the action actually started if the action was carried out simultaneous to the participants’ talk. After repeatedly viewing the recorded data, I began to notice how often the participants within the sandpit carried out activities in the sand, while simultaneously conducting talk. This almost dual-interaction, between the participants and the sandpit, cased me to rethink the way I was representing the data (thesis S, p. 155)

In these words we see the researcher pushing the boundaries of the Jefferson notation system in order to meet the analytic needs of her own study. This is clearly stated in sections of the thesis and explained and illustrated. She also acknowledges drawing on the work of other researchers who use different transcription methods such as Ochs (1979) and Bloom (1993). In this way, she shows the intellectual work she is engaged in, in relation to transcription and transcription as an aspect of a systematic endeavour to examine a problem. Would we expect this level of detail in studies informed by approaches such as grounded theory? Perhaps not, however, we suggest that encompassing decisions made about transcription especially in relation to issues and problems encountered, could be a useful thing to do, and a way to strengthen the “show” of understanding

about research that is integral to the production of a thesis. We found the researcher's work on transcription, not only informative but useful for our work as researchers who develop transcripts and examine interviews and classroom interactions.

There are a few studies in the literature of challenges that face researchers when transcribing, and some studies that give valuable insight into the transcription process. However, it may be necessary to step outside the canon that informs our own particular research approach and to embrace more transcription literature that informs qualitative inquiry in general, and qualitative methods specifically.

Reporting transcription as an aspect of study design and research logic

As the analysis progressed, we noted the use of Crotty's (1998) work and its influence on perspectives that informed the Australian doctoral work. So, while words such as 'epistemology' and 'ontology' were not a focus for our content analysis, these words were noticed as they recurred across a number of theses. Later, it seemed to us that the use of theoretical perspectives within qualitative inquiry was related to the use of specific research approaches and methods but was not necessarily considered as connected to transcription and use of transcripts. That is, while researchers acknowledged the constructed nature of reality, for example, or the interpretive work of research participants in research, these same theoretical understandings and everyday practices were not related to transcription in explicit ways. That is, the lack of comment about transcription as selective and partial, for example, could be taken to infer that transcripts were not selective and partial. In the literature, Lindsay (2000) refers to a similar absence noted by her in higher degrees assignments as 'positivism by default' and it is on the basis of this that she calls for researchers to acknowledge the partial and selective nature of transcription and transcripts – or at least, to make a position explicit in the reporting of research. Our observation of the noticeable absence of links between theoretical perspectives on knowledge and reality and transcription, in qualitative research, support Lindsay's findings that many qualitative researchers do not comment on transcription in relation to theoretical perspectives. Further, we tentatively suggest that making this link would enhance the reporting of studies, particularly in relation to how trustworthiness is addressed.

Conclusions

There are numerous claims made in the literature about the ways that transcription is taken for granted in qualitative inquiry. Few of these claims have been substantiated through empirical studies of transcription. In this paper we have outlined and considered a pilot study of accounts of transcription provided by a small number of doctoral students in their theses. We conclude that some aspects of transcription are taken for granted in the reporting of PhD research. However, this taken-for-grantedness appears related to the various methodological perspectives that inform the work of researchers, not necessarily the "oversights" of individuals. Common patterns within the four methodologies we examined perhaps reveal more about transcription as a practice than patterns across the corpus. At the same time, we suggest that shared omissions across the theses, together with ways that some individual researchers addressed them, reveal fruitful directions that many qualitative researchers in education can draw from in the reporting of transcription. Specifically, use of the transcription literature, acknowledgement of how transcription was conducted and addressing transcription as an aspect of trustworthiness are all worthwhile practices. The matter of a notation system is a more challenging aspect. Certainly,

we would not endorse the position that all researchers should use standardized and specific systems – at the same time, the use of modified systems or literacized practices needs to be carefully explained and documented in the thesis. The study suggests the need for greater attention to transcription in the reporting of doctoral research, and for more empirical studies to examine this aspect of reporting of qualitative research especially as it is addressed within particular methodologies.

References

- Atkinson, J. M., & Heritage, J. (1999). Jefferson's transcript notation. In A. Jaworski & N. Coupland (Eds.), *The discourse reader* (pp. 158-166). London; New York: Routledge.
- Davidson. (2009). Transcription: Imperative for qualitative research. *International Journal of Qualitative Methods*, 8(2), 35-62.
- Baker, C. (1997). Transcription and representation in literacy research. In J. Flood, S. B. Heath & D. Lapp (Eds.), *Handbook of research on teaching literacy through the communicative and visual arts* (pp. 110-120). London: Prentice Hall Interactional.
- Berg, B. L. (2004). *Qualitative Research Methods for the Social Sciences* (5th ed.). Boston, MA: Pearson Education.
- Bird, C. M. (2005). How I stopped dreading and learned to love transcription. *Qualitative Inquiry*, 11(2), 226-248.
- Bloom, L. (1993). *Transcription and coding for child language research: The parts are more than the whole*. In J. A. Edwards & M. D. Lampert (Eds.), *Talking data: Transcribing and coding in discourse research* (pp. 149 – 166). Hillsdale, NJ: Lawrence Erlbaum.
- Bucholtz, M. (2000). The politics of transcription. *Journal of Pragmatics*, 32(2000), 1439-1465.
- Coates, J., & Thornborrow, J. (1999). Myths, lies and audiotapes: Some thoughts on data transcripts. *Discourse and Society*, 10(4), 594-597.
- Cook, G. (1990). Transcribing infinity. *Journal of Pragmatics*, 14, 1-24.
- Crotty, M. (1998). *The foundations of social research: Meanings and perspectives in the research process*. Crows Nest, NSW: Allen & Unwin.
- Du Bois, J. W. (1991). Transcription design principles for spoken discourse research. *Pragmatics*, 1(1), 71 - 106.
- Duranti, A. (1997). *Linguistic anthropology*. Cambridge: Cambridge University Press.
- Duranti, A. (2007). Transcripts, like shadows on a wall. *Mind, Culture and Activity*, 13(4), 301-310.
- 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.
- Feeley, N., & Gottlieb, L. N. (1998). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nursing Education Today*, 24(2), 105-112.
- Gomm, R. (2004). *Social research methodology: A critical introduction*. Basingstoke: Palgrave Macmillan.
- Graneheim, U. H., & Lundman B. (2004). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nursing Education Today* 24,105-112.

- Green, J., Franquiz, M., & Dixon, C. (1997). The myth of the objective transcript: Transcribing as a situated act. *TESOL Quarterly*, 21(1), 172-176.
- Grundy, A. L., Pollon, D. E., & McGinn, M. (2003). The participant as transcriptionist: Methodological advantages of a collaborative and inclusive research practice. *International Journal of Qualitative Methods*, 2(2). Retrieved September 23, 2008 http://www.ualberta.ca/~iiqm/backissues/2_2/html/grundyetal.htm
- Halcomb, E. J., & Davidson, P. M. (2006). Is verbatim transcription of interview data always necessary? *Applied Nursing Research*, 19, 38-42.
- Hepburn, A. (2004). Crying: Notes on description, transcription, and interaction. *Research on Language and Social Interaction*, 37(3), 251-290.
- Jaffe, A. (2007). Variability in transcription and the complexities of representation, authority, and voice. *Discourse Studies* 9(6), 837-842.
- Jaffe, A. (2000). Introduction: Non-standard orthography and non-standard speech. *Journal of Sociolinguistics*, 4(4), 497-513.
- Kohlbacker, F. (2006). *The use of qualitative content analysis in case study research*, 7(1). Retrieved 20 July, 2009 <http://www.qualitative-research.net/index.php/fqs/article/viewArticle/75>
- Kvale, S. (1996). *Interviews: An introduction to qualitative research interviewing*. Thousand Oaks, CA: Sage.
- Lankshear, C., & Knobel, M. (2004). *A handbook for teacher research: From design to implementation*. Maidenhead, Berkshire: Open University Press.
- Lapadat, J. C. (2000). Problematising transcription: Purpose, paradigm and quality. *Social Research Methodology*, 3(3), 203-219.
- Lapadat, J. C., & Lindsay, A. C. (1998, April). *Examining transcription: A theory-laden methodology*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.
- Lapadat, J., & Lindsay, A. C. (1999). Transcription in research and practice: From standardization of technique to interpretive positionings. *Qualitative Inquiry*, 5(1), 64-86.
- Luebs, M. (1996). *Frozen speech: The rhetoric of transcription*. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.
- MacLean, L. M., Meyer, M., & Estable, A. (2004). Improving accuracy of transcripts in qualitative research. *Qualitative Health Research*, 14(1), 113-123.
- MacWhinney, B., & Snow, C. (1991). The wheat and the chaff or four confusions regarding CHILDES. *Journal of Child Language*, 19, 435-458.
- Mertens, D. E. (1998). *Research methods in education and psychology: Integrating diversity with quantitative and qualitative approaches*. Thousand Oaks, CA; London; New Delhi: SAGE.
- Mischler, E. (1991). Representing discourse: The rhetoric of transcription. *Journal of Narrative and Life History*, 1(4), 255-280.
- Mondada, L. (2007). Commentary: Transcript variations and the indexicality of transcribing practices. *Discourse Studies*, 9(6), 809-821.
- Ochs, E. (1979). Transcription as theory. In E. Ochs and B. B Schiefflin (Eds.), *Developmental pragmatics* (pp. 43-72). New York: Academic.
- O'Connell, D. C., & Kowal, S. (1999). Transcription and the issue of standardization. *Journal of Psycholinguistic Research*, 28(2), 103-120.

- 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.
- Poland, B. D. (1995). Transcription quality as an aspect of rigor in qualitative research. *Qualitative Inquiry*, 1(3), 290-310.
- Psathas, G. (1995). *Conversation analysis: The study of talk-in-interaction*. Thousand Oaks, CA: Sage.
- Rossmann, G. B., & Rallis, S. F. (2003). *Learning in the field: An introduction to qualitative research* (2nd ed.). Thousand Oaks, CA; London; New Delhi: SAGE.
- Sacks, H. (1995) *Lectures on Conversation/Harvey Sacks; edited by Gail Jefferson; with an Introduction by Emanuel A. Schegloff*. Oxford: Blackwell.
- ten Have, P. (2007). *Doing Conversation Analysis: A practical guide 2nd edition*. London: SAGE Publications.
- Tilley, S. A. (2003a). Challenging research practices: Turning a critical lens on the work of transcription. *Qualitative Inquiry*, 9(5), 750-773.
- Tilley, S. A. (2003b). Transcription work: Learning through coparticipation in research practices. *Qualitative Studies in Education*, 16(6), 835-851.
- Tilley, S. A., & Powick, K. D. (2002). Distanced data: Transcribing other people's research tapes. *Canadian Journal of Education*, 27(2/3), 291-310.
- Wellard, S., & McKenna, L. (2001). Turning tapes into text: Issues surrounding the transcription of interviews. *Contemporary Nurse*, 11, 180-186.
- Wilkinson, S. (2004). Focus groups. In David Silverman (Ed), *Qualitative research: theory, method and practice*, (pp. 177-199). Thousand Oaks: Sage Publications.