SIX
Transcribing Audio and Video Data

Reflexive practice and ethics	Analysing textual data	Analysing image, audio, and video data
Collaborating	Transcribing	Representing findings
Reviewing literature	Generating data	

Companion website materials available here:

www.uk.sagepub.com/paulus

Learning Objectives

- Reflect upon the ways in which transcription is a situated act.
- Compare and contrast four approaches to transcription (verbatim, Jeffersonian, gisting and visual).
- Select and critique appropriate digital tools for transcribing audio and video data.
- Have strategies to ensure ethical practices when transcribing data.

Introduction

In Chapter 5, we discussed the various ways in which digital tools can be used for collecting both naturally occurring and researcher-generated data, such as recording conversations and interviews in new ways. In this chapter, we turn to

digital tools that can support transforming recorded data into transcriptions for further analysis. We discuss how transcription itself is a situated act, describe four different types of transcription (verbatim, Jeffersonian, gisted and visual) and discuss features of four software packages (Express Scribe, Audio Notetaker, InqScribe and F4/F5). We also explore using voice recognition software (Dragon Dictate) and the video analysis tool Transana.

Transcription as Situated Act

Today's recording devices allow us to record vast amounts of audio and video data that must somehow be made sense of. Instead of being limited by the amount of information that could be documented by hand, we are now limited by how much recorded information we can process (Lee, 2004). As Ochs (1979) suggested, simply recording everything without making decisions about what is important results in these decisions being pushed to the transcription phase.

There are many approaches to transcription, depending on your field and understanding of what the transcript represents. For example, if you are interested in the way government officials' language is gendered you may be collecting, recording and transcribing data very differently than if you are looking at the use of turn-taking in conversations amongst doctors and their receptionists. What and how you choose to transcribe should be closely connected to your research focus and methodological approach, resulting in certain types of transcripts.

As you transcribe, you make particular choices, and those choices are related to your theoretical stance (Kvale, 2007). When creating a verbatim (word-for-word) transcript, for example, you make frequent decisions such as whether and how to include informal speech, non-word utterances, repetitions, stuttering, interruptions and background or other incidental sounds; see Figure 6.1.

Helen:	count your cards
Brock:	five six
Joseph:	I have six
Helen:	two three four five
Joseph:	I have six
Helen:	oh put it back on the pile then thank you
Joseph:	one two three four five
Brock:	one two three four five I have five
Helen:	awesome okay Joseph since you got to deal the cards your job is to pick a friend to go first
Joseph:	Helen
Helen:	Oh thank you all right it's your job first look at your cards and see if you have what colour

Figure 6.1   Example of a verbatim transcript

The interaction in Figure 6.1 included three young boys (Joseph, Brock and Luke) and their therapist (Helen). However, only three speakers (Helen, Joseph and Brock) are represented in the transcript, as Luke was non-verbal and used body language to communicate. Choosing to only transcribe spoken words, and exclude physical movement or non-verbal behaviour, flattens the data and has consequences for how the interaction is being represented. Because it is impossible to document all features of social interaction, all transcripts should be considered *partial* representations, *selective* and *situated* in relationship to the goals of a particular study (Davidson, 2009; Lapadat and Lindsay, 1999).

Because transcription can be approached from a variety of perspectives, it is important that you engage in reflexive practice around your transcription choices. One view is that transcription can be seen as a form of analysis itself (e.g. Hammersley, 2010). Another is that qualitative data analysis software is making full transcription obsolete (Evers, 2011). Those who don't think about their approach in advance may end up with transcripts that do not align with their research goals (Oliver, Serovich and Mason, 2005). Thus, we encourage you to take a moment to complete the Reflexive Practice 6.1 activity.

REFLEXIVE PRACTICE 6.1

Consider the study you are designing in conjunction with this book:

- What kind of audio or video data might you collect and transcribe?
- What information in the recording do you need to attend to?
- Will features such as rate of speech, pitch and/or overlapping speech be important?
- How might you capture such features in your transcript?
- How might you represent laughter, pauses and gestures?

Transcription requires an investment of time, money and other resources. It takes a minimum of four hours to create a verbatim transcript for one hour of a high-quality recording (Dempster and Woods, 2011; Evers, 2011). Recordings of poor quality or of multiple overlapping speakers will take longer to transcribe. Case Study 6.1 illustrates some factors to consider in deciding on a transcription approach.

CASE STUDY 6.1

Jimmy is interested in interviewing nurses that have left the profession. He is considering conducting and audio-recording in-depth interviews, lasting about two hours, with 25 or more nurses, ex-nurses and other health workers. He realizes that it will take him around six hours to transcribe a one-hour interview. Based on this calculation, he would be under-taking 300 hours of transcription. After thinking about the resources he has, he decides to redesign his study to ask some of the questions in an online form that will allow participants

(Continued)

> *(Continued)*
>
> to respond at length, reducing the number of questions to be asked in the interview itself. He can then compare the participants' online responses with their interview transcripts, analysing across the data. Further, he decides not to transcribe the audio files in their entirety. Rather, he uploads his audio files into a software program that allows him to directly code and memo the audio files as well as transcribe selected portions.

We next discuss several ways to transcribe, each of which represents the data in a unique way. First, consider Reflexive Practice 6.2.

REFLEXIVE PRACTICE 6.2

Replacing real names with pseudonyms is standard practice in qualitative research. Deciding when and how to remove this identifying information often poses a dilemma for researchers. How might digital tools make it easier, yet more difficult, to anonymize your data? When will you apply pseudonyms – during or after transcription?

Types of Transcription

In this section, we discuss four types of transcription: verbatim (representing word-for-word what is said), Jeffersonian (representing additional features of the talk beyond the words), gisting (representing just an essence or condensed version) and visual (representing meaning with images).

Verbatim Transcription

A verbatim transcription involves typing everything you hear (in an audio recording) and/or see (in a video recording). This includes representing all utterances made by all participants without changing non-standard language usage (e.g. 'he don't care about me') or dialect (e.g. Doric Scots usage of 'fit like?' for 'how are you?') and without skipping over repetitions ('and-, and-'), false starts ('uh-, well, I mean') and backchannels ('mm-hmm'). Creating a verbatim transcript from video data should also include all of the non-verbal communicative behaviours (e.g. yawning, raising hands, throwing hands up in the air). While in theory a verbatim transcript is one that captures 'everything', it is never really possible to capture all that is communicated. To capture as much as possible, however, you will need to engage in several cycles of transcription, reviewing the recording multiple times. Foot pedals and shortcut keys can help with this.

Jeffersonian Transcription

Anyone who has attempted to create a verbatim transcript will have encountered the difficulty of representing features of the talk such as the rate of speech, volume and overlapping speech. In research traditions such as conversation analysis these features are assumed to carry meaning and so it is important to

note them. There is a variety of notation systems (e.g. Du Bois, 1991; Gumperz and Berenz, 1993) that provide a way to represent these features. We focus here on Jefferson's (2004) notation system (see Figure 6.2 for a modified version of the notation system and Figure 6.3 for a modified Jeffersonian transcript).

↑	Upward arrows represent marked rise in pitch.
↓	Downward arrows represent a downward shift in pitch.
> <	Text encased in 'greater than' and 'less than' symbols is hearable as faster than the surrounding speech.
=	Equal signs at the end of a speaker's utterance and at the start of the next utterance represent the absence of a discernible gap.
the	Underlining represents a sound or word(s) uttered with added emphasis.
[ ]	Extended square brackets mark overlap between utterances.
(7)	Numbers in parentheses indicate pauses timed to the nearest second. A period with no number following, (.), indicates a pause which is hearable, yet too short to measure.

Figure 6.2   Modified Jeffersonian (2004) notation system

1   *Bria*: Okay so (.) what do you think on a scale of one to ten↑ is (.) the whole
2   grace situation for you at summer school right now↑ (.)
3   *Devin*: Well like (.) with one being handling it good or handling it bad↑
4   *Bria*: Uh (.) let's make one handling it bad and ten (.) handling it excellently↓ (2)
5   *Devin*: Well then it's probably a (.) somewhere between two and four↑=
6   *Bria*: Yeah that's that was kind of my impression in fact I'm not=
7   *Devin*: =Somewhere between two and four and a half probably↓ (.)

Figure 6.3   Excerpt from a modified Jeffersonian transcript from Lester and Paulus (in press)

Visit Web Resource 6.1 for more on transcription notation systems.

Jeffersonian transcripts must be created in rounds (ten Have, 2007), in which the transcriber focuses on different features of the talk each time. For instance, in the first round you might focus on the pauses between conversational turns, while during the second you might focus on intonation. Overall, this transcription approach involves multiple listenings, identifying analytically interesting sections for in-depth transcription, and using notation symbols to capture features which would not be included in other transcript types.

Case Study 6.2 illustrates how a researcher moved from hundreds of hours of video data to focused Jeffersonian transcripts. Further, the researcher decided to use Transana, a software program designed specifically for video transcription and analysis, because it supports Jeffersonian notations. Transana is discussed later in this chapter and further in Chapter 8.

CASE STUDY 6.2

Abraham has a corpus of video data from a youth therapy centre. Each therapy session lasted 30 minutes. Prior to transcribing his data, he listened to each recording once to become familiar with it. After completing one round of verbatim transcription with approximately 25 hours of data, he listened to the therapy sessions again and refined the transcriptions. As he gained a deeper level of familiarity with the data set, he began to attend only to those sections that focused on 'dealing with reported problems' (e.g. hitting a teacher). He used Transana to pull out just those sections and transcribed them using Jeffersonian notations. Transana was ideal for this as he was able to use the symbols already included within the software while listening and re-listening to capture detailed features of the talk such as interruptions and overlapping speech. He could easily take a few seconds of the talk and listen repeatedly to interpret where the overlapping speech occurred.

Visit Web Resource 6.2 for a Jeffersonian transcription tutorial.

Gisted Transcription

On the other end of the spectrum are researchers for whom gisting may be a viable option. A gisted transcript is similar to news show reports sharing the highlights of a politician's speech at the Houses of Parliament. Figure 6.4 defines the two types of gisted transcript: condensed and essence.

The process of creating a condensed transcript involves listening to the recording and leaving out 'all the utterings which do not seem relevant to the research question' (Evers, 2011, p. 13). All backchannels ('umm', 'er') are left out. Here is an example (the dots '…' indicate omitted speech):

> 90% of my communication is with … the sales director. 1% of his communication … with me. I try to be one step ahead, I get things ready … because he jumps from one … project to another. This morning we did Essex, this afternoon we did BT. (Evers, 2011, p. 13)

One of the most challenging issues with the condensed transcript is deciding what to leave out, while still retaining enough context for analytic purposes.

The second type of gisted transcript is the essence transcript (Dempster and Woods, 2011). While the condensed transcript captures the exact words, the

The condensed transcript	The transcript is condensed by removing unnecessary words and phrases, leaving a simplified version but with exact words. No additional text is added.
The essence transcript	The transcript retains the essence of the event through paraphrasing. It may even have single-word sections. Often used in tandem with software which hyperlinks back to the original media file. It can also include pictures in the transcript (visual transcript) instead of or in addition to words.

Figure 6.4 Types of gisted transcript

essence transcript retains only a paraphrased version of the recorded data. Dempster and Woods (2011) described this type of gisting as creating:

> a summary transcript that captures the essence of a media file's content without taking the same amount of time or resources as a verbatim transcript might require. Typically, a transcriber … may take four or five hours to create a verbatim transcript of the spoken word in a typical hour-long media file, while such a file can be gisted in one to two hours. (p. 22)

For instance, say that data was collected from a mathematics class in which children are studying patterns, similarity and symmetry. Some used actual quilt square patterns and others used a computer simulation. Figures 6.5 and 6.6 (where 'T' is the teacher, 'S' is the student) illustrate the differences between a verbatim transcript and an essence transcript.

Six turns in the verbatim transcript are represented as simply 'alternate explanations' in the essence transcript. Essence transcripts rely on your ability to adequately summarize the data based on your research purpose. This type of approach means that a heavier layer of interpretation is occurring whilst transcribing. What is kept in and left out becomes a more overt analytical act.

Using a tool such as InqScribe or Transana can be useful for gisted transcripts. These tools allow you to synchronize the transcripts with the media file. In this way, portions of the full recording can be revisited later if they become important. This means that you can be much more selective with what you transcribe, without losing access to other parts of the data.

T: Why do you think it is that the computer with the quilting software, since it's, since that software is for us to use to design quilts, why won't it let us do something we found out about quilts, which is diagonal flip?

S: Maybe they've never [heard of them.]

S: [Maybe the turtle…]

T: Maybe they've never heard of them.

S: Maybe the [turtle…]

T: [Do you think] that's possible?

S: Yeah.

S: [Maybe there is no, maybe there…]

S: [Maybe the people didn't, just didn't know] Maybe the people that, um, put the, um, the, um, the thing in didn't know about diagonal flips yet.

S: Or maybe, like, there wasn't any, like, CC, like C…G, or like, or any code thing.

T: Any command? There's not a [command.]

S: [Like D.]

S: Maybe they didn't get the right computer chip for that.

Figure 6.5 Verbatim transcript

T: Why doesn't computer know diagonal flips?

S: Maybe they've never heard of them.

T: Do you think that's possible?

S: Yeah.

S: Alternate explanations

Figure 6.6 Essence transcript

Visual Transcription

Essence transcripts will often contain images to represent the action. This kind of 'visual transcription' uses still images taken from a video recording to represent the meaning. In Figure 6.7, still images taken from a video are combined with the descriptions in the transcript.

Synchronizing the media file with the transcript, you can click on the image to go directly to that point in the video file. This makes it possible to 'bookmark'
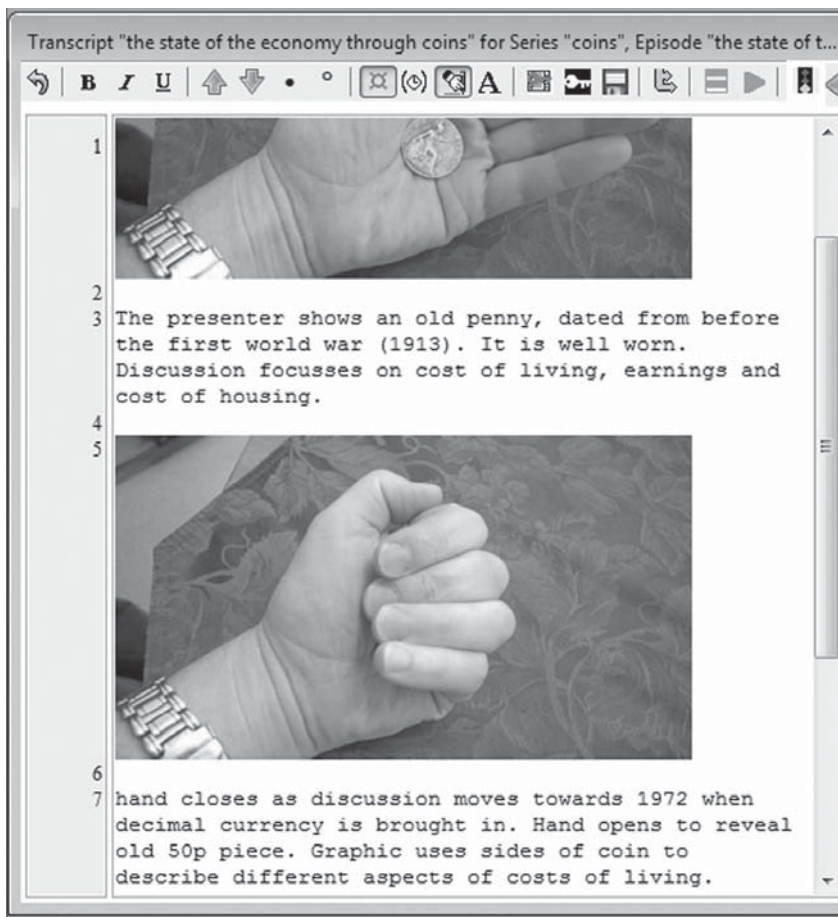

Figure 6.7 Visual transcript

hours of video and stay close to the original data. This approach, described more fully in the last section of this chapter, allows the researcher to analyse the audible layer, the visual layer and the interaction between these two layers (Dempster and Woods, 2011). See Reflexive Practice 6.3.

REFLEXIVE PRACTICE 6.3

Which of the transcription types do you have experience with? What type of transcription will be useful in the study you are designing?

Digital Tools to Support the Transcription Process

The development of recording technologies has impacted qualitative research in ways that few could have foreseen. The advent of reproducible sound recordings in the late nineteenth century did not mark the end of written field notes, as sound recordings remained a fragile technology. Indeed, researchers often made use of professional note takers (stenographers), which limited where interviews could be conducted. For example, the Chicago School researchers would access bureaucratically defined populations, such as prisoners, and interview them in office locations. 'The use of stenographers presumably fitted rather easily into this kind of "fieldwork of the office." While stenography involved a shift towards real-time recording, it was inevitably spatially constrained' (Lee, 2004, p. 874). Later, compact cassette tapes became popular devices for recording interviews and dictations. Specialized transcribing machines and foot pedals made the process even more efficient.
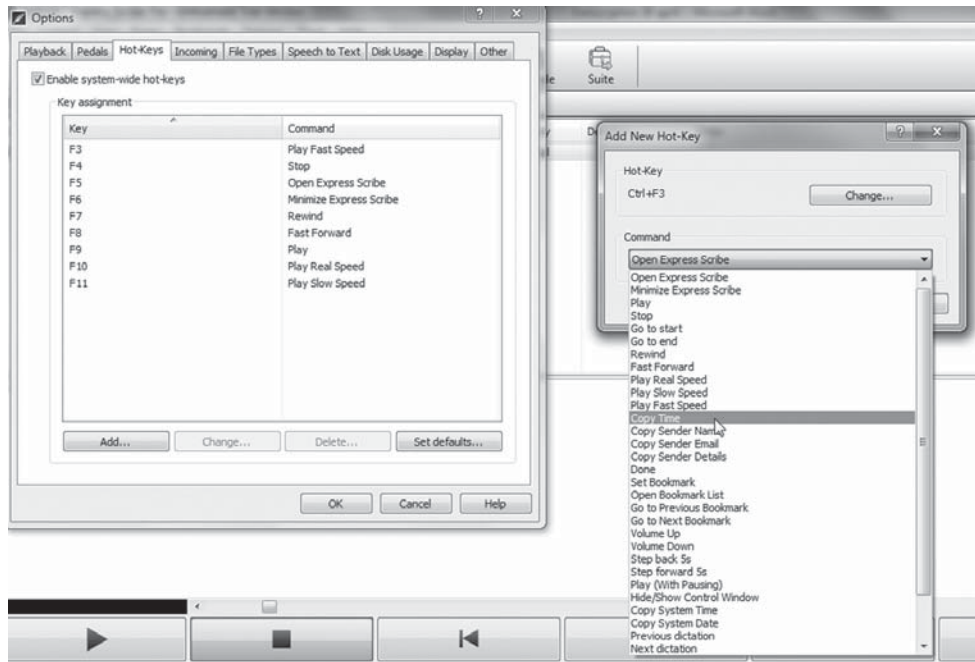
Kelle (2007) noted that the rise of informal qualitative interviewing coincided with the technological development of the portable tape recorder, transforming the 'non-recurring encounter' (p. 444) between the participants and researcher into an event that could be re-examined numerous times by the researcher (p. 444). With tape recorders, interviewing techniques such as focus groups became easier to manage. More diverse examples of data collection became possible, such as the 'walk-along' (Pink, 2007; 2009), described in Chapter 5, and soundscapes (Hall, Lashua and Coffey, 2008), described in Chapter 8. Recording, then, has moved qualitative research well beyond the bureaucratic spaces of the Chicago School into everyday life.

We discuss next four transcription software packages: Express Scribe, Audio Notetaker, InqScribe and F4/F5.

Express Scribe by NCH Software

Express Scribe, which works alongside a word processor, is a cost-effective and simple program. It works with both Windows and Apple operating systems and

provides keyboard shortcut 'hot keys' to stop, start, rewind and fast forward the recording (see Figure 6.8). The professional version supports the use of foot pedals, templates and email file exchange. It does not, however, provide a way to visually conceptualize your data or synchronize the media file with the transcript.

Figure 6.8 Express Scribe shortcut keys

Visit Web Resource 6.3 to watch a screencast introduction to Express Scribe.

Sonocent™ Audio Notetaker 3

Audio Notetaker supports audio data and works with both Windows and Apple operating systems. Audio Notetaker also works well in combination with voice recognition software, such as Dragon Dictate (described later in this chapter). A novel feature is its visual representation of the audio file as colour-coded blocks. This allows you to identify key parts of the recording and, with a click, return to that section of the data. Clicking on the colour-coded bars allows you to copy, paste or delete sections of the recording. Another useful feature is its ability to split the audio into sections and highlight blocks in different colours to mark analytically important speech that needs to be transcribed. This can also be done while recording live on your computer, which can be useful if you want to mark important parts of the audio data in real time. Unfortunately, these 'blocks' are not linked to the final transcript. In Figure 6.9 the blocks on the right link directly to places in the media file.

Figure 6.9 Audio blocks in Audio Notetaker 3 transcript

Visit Web Resource 6.4 to read case studies of researchers using Audio Notetaker.

Audio Notetaker transcripts are searchable, and you can export them along with the audio file to your portable computing device and/or iTunes. Cleverly, the audio file can be exported as an album, with the notes that you have taken becoming the track name and any images that you have included becoming the album art. Overall, the software provides a number of different ways to organize your media files visually and textually, making the transcript somewhat interactive and speeding up the transcription process. Case Study 6.3 illustrates this.

CASE STUDY 6.3

Dorothy is an ethnographer, working on a busy hospital ward. Unable to take extensive written field notes, she uses a recorder to dictate her initial notes, elaborating upon them later when back in her office. She uses a 'debrief file', which involves her talking to the recorder on her commute home. These musings are loaded into Audio Notetaker. When later listening to the recordings, long pauses, such as when she was concentrating on the road, can be quickly skipped by looking at the visual representation of the media file (the blocks). Once she has finished transcribing the notes these are imported into a word processor.

Inquiriam's InqScribe 2.1.1

InqScribe is transcription and subtitling software that works with both audio and video data and is compatible with both Windows and Apple operating systems. Unlike Express Scribe, which works with a word processor to create the transcript, with InqScribe you transcribe directly into the provided window using media file playback controls. Its key features include shortcut keys to control playback, the ability to create macros ('snippets') and synchronization of the media file with the transcript.

Synchronizing the media file with the transcript can lead to a more transparent, rigorous study, as it makes it possible to review the original data recording rather than rely solely on the transcript (Ashmore and Reed, 2000). The original source of data can be made available not only to all members of the research team, but also to the participants and even the reader, providing new opportunities for member checking and strengthening the trustworthiness of the study. This may also, however, raise some ethical concerns, which we ask you to consider in Reflexive Practice 6.4.

—————— **REFLEXIVE PRACTICE 6.4** ——————

What ethical dilemmas might be raised when synchronizing the media files with the transcript and then making them available to participants and/or consumers of your research findings?

Table 6.1 outlines some examples of how a synchronized transcript may be useful.

Table 6.1   Uses of synchronized transcripts

Synchronized transcripts can …
… make it easier to listen again to particular points in the recording, rather than relying only on the transcript during analysis.
… make it easier to navigate from one section of a long transcript to another.
… be useful for reviewing the full portion of a media file that was gisted.
… serve as a point of reference for other researchers in the team who were not involved in transcription.

In Vignette 6.1, David Woods of Transana discusses how moving seamlessly between gisted transcripts and the original media allows researchers to stay close to their data.

Vignette 6.1 ⊣ **Transcription, Transcripts and Software**

David Woods, University of Wisconsin-Madison, Developer of Transana

Recorded data can be described as a selected abstraction of reality, and transcription is a further selected abstraction of the recording. Further, analysis is a selected abstraction of the transcript, and writing up your findings is often yet another selected abstraction of the analysis. This successive step-wise abstraction (also known as layers of interpretation) is often necessary in the process of moving from raw data to theory, but often comes at the cost of increasing distance between the researcher and the original data. The challenge for the researcher is to be careful to remain 'close' to the data at all stages of the analytic process. This can be accomplished rather well in the transcription and transcript review phases of analysis by using software that links and synchronizes the original media file data with the transcript, allowing the researcher to hear (and in the case of video, to see) the data being referred to while looking at the associated transcript. The written representation of the data (the transcript), is useful for many analytic tasks, but there is no true substitute for listening to the inflection and tone of how the words were said, or seeing the facial expression and body language of the speaker. Having the transcript linked to the media file gives the researcher significantly more analytic and interpretive power than working from the transcript alone.

Synchronizing transcripts in InqScribe is quite simple with the use of time stamps. When the media file is playing, or paused at a point where you want to add a time

code, you simply press the Ctrl-+ shortcut which inserts the time into the transcript. Figure 6.10 shows time stamps at the end of each speaker's turn (e.g. [00:00:29:22]) Clicking on the time stamp replays the recording at that point in the transcript.
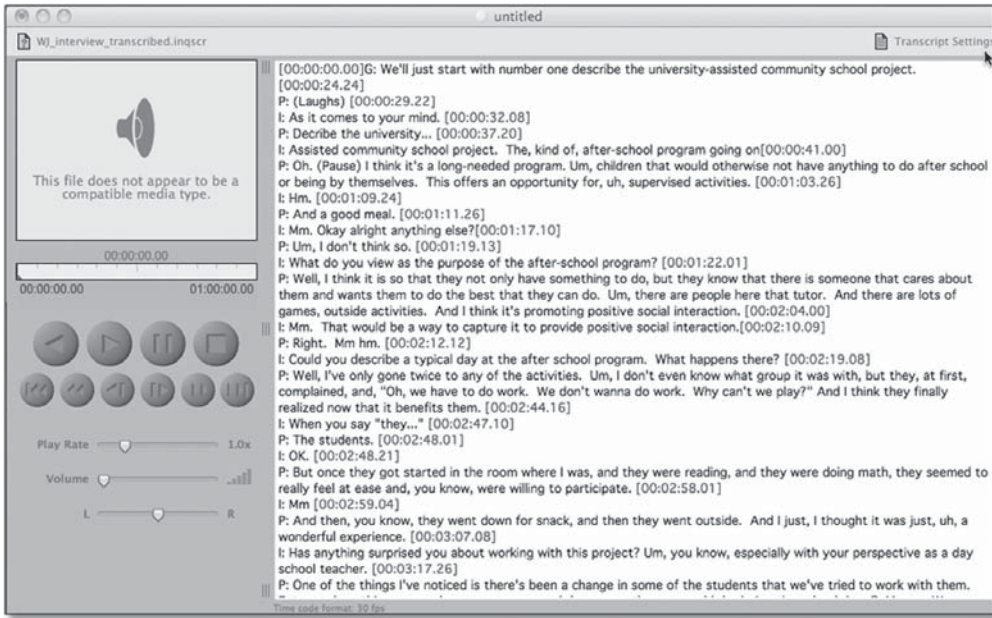

Figure 6.10 Transcript with time stamps in InqScribe

InqScribe also supports the use of shortcuts and customizable macros called 'snippets' to make transcribing easier (see Figure 6.11). In Vignette 6.2, Joshua Johnston discusses his use of snippets to support transcription.
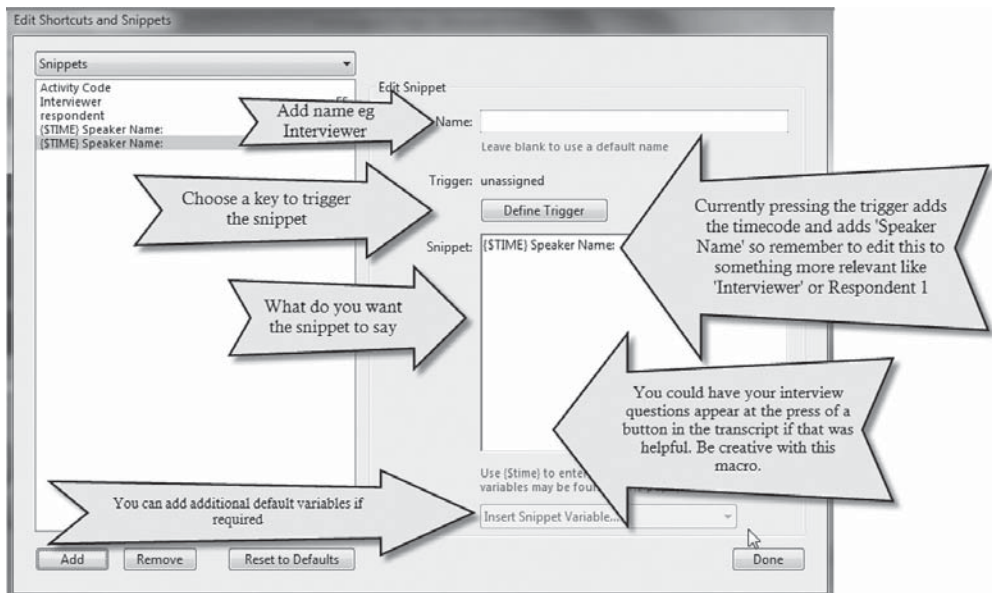

Figure 6.11 Creating snippets in InqScribe

Vignette 6.2	**Creating and Using Customizable Shortcuts with InqScribe**

Joshua P. Johnston, University of Tennessee

As a discourse analyst, I rely on InqScribe to prepare working drafts of my transcripts. What makes the program especially helpful are the 'snippets' – key combinations that insert commonly used words or phrases.
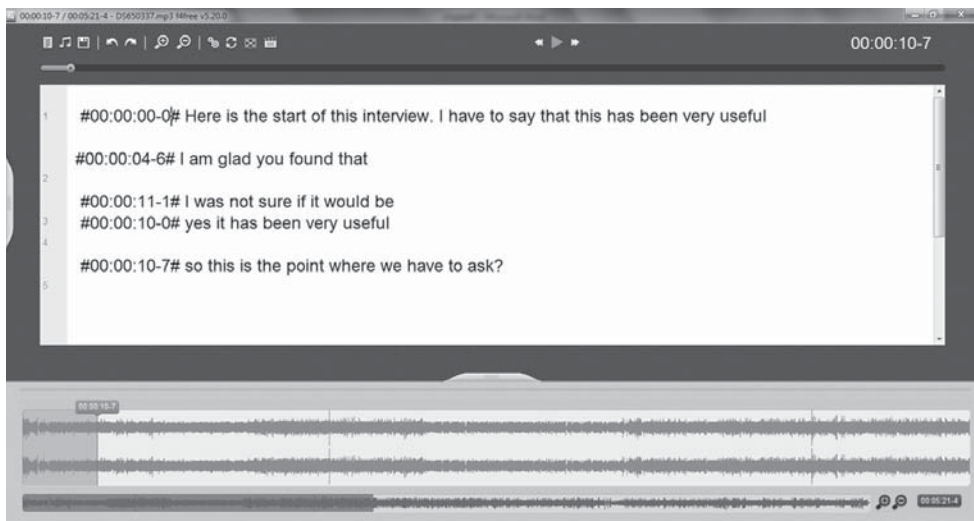
I use snippets in two ways to speed up my transcribing. First, I assign commonly used phrases and verbal tics of the participants to keys that I don't use otherwise. I assign [0] to 'umm', [-] to 'like' and [=] to 'yeah'. Second, I use snippets to aid in the formatting of my transcript. For example, I create snippets to alternate among speaker names and to insert a double space before participants' pseudonyms with the push of a single button. I create snippets named for each participant and assign each one to a [Function] key. Once defined, I simply press the corresponding [Function] key when I come to the end of one participant's utterance and the beginning of another's.

Understandably, you may feel like remembering all the snippets would be unmanageable, but you can keep the names and triggers visible in a box at the side of your screen. The snippets serve to maintain confidentiality by inserting pseudonyms rather than real names into the transcript.

F4 (Windows Version) / F5 (Apple Version)

F4/F5 supports audio and video data. It also supports adding time stamps to the data, as shown in Figure 6.12 (e.g. #00:00:04-6#).

Shortcut keys in F4/F5 can be used to insert the name of the speaker automatically at the beginning of a new paragraph and to support Jeffersonian or

Figure 6.12 F4/F5 transcription software main screen

other notation systems. One of the greatest benefits of this software is that the transcripts will load straight into the qualitative data analysis software packages MAXQDA and ATLAS.ti without losing the notation symbols or time codes.

One other approach to transcription is to directly type into one of the major data analysis software packages, described in Chapter 7, which also support transcription and synchronization of the media files. In Vignette 6.3 Susanne Friese describes the use of ATLAS.ti for doing so.

Vignette 6.3 | **Associated Documents in ATLAS.ti**

Susanne Friese, Support Specialist, ATLAS.ti

The associated documents function allows you to transcribe your audio or video data and link documents to each other so that you can view them synchronously. This means you can mark a segment or a quotation in an associated document and instantly listen to the original audio or view the video.

As shown in the figure below, you can highlight a piece of text in the associated file and press the key combination Ctrl-P to play the linked video segment. The associated recording will be played back from the correct position. Playback stops at the end of the selection. You can also use the association points as a navigation device, either to enter at a particular point in the file or to easily jump around.
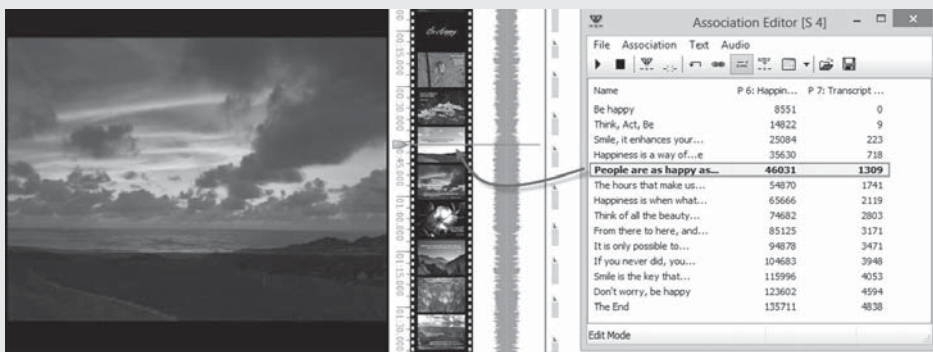
Table 6.2 provides a summary of features across the four transcription software packages that have been discussed so far.

Table 6.2 Transcription software features compared

Express Scribe	• Free version (does not expire) • Mac/PC platform • Supports audio and video data • Fee-based version supports foot pedal use • Hotkeys to control playback • Integrates with word processing programs • Works with voice recognition software
Audio Notetaker 3	• 30-day fully functional free trial version; offers site licences • Mac/PC platform • Supports audio data only, but can import PowerPoint slides • Supports foot pedal use • Audio represented visually with coloured bar segments for navigation, editing and organizing • Works well with voice recognition software
InqScribe 2.2.1	• 30-day fully functional free trial version with academic discounts • Mac/PC platform • Supports audio and video data • Supports foot pedal use • Can insert time codes for transcript/media file synchronization • Can create macros (snippets) for repeated words and phrases • Works with voice recognition software
F4 / F5	• Free version very limited • Mac/PC platform • Supports audio and video data • Supports foot pedal use (in pro version) • Hotkeys to control playback, add timestamps and other shortcuts • Integrates with word processing programs and with CAQDAS tools • Does not work well with voice recognition software

Now take a look at Reflexive Practice 6.5.

───── **REFLEXIVE PRACTICE 6.5** ─────

Take some of your recorded data from the Chapter 5 reflexive practice activities, choose a transcription approach and a transcription tool. Practise transcribing your file. Reflect on your experience researching and testing out the selected transcription tool.

Other Approaches to Transcription

We next explore two other approaches to transcribing – using voice recognition software (such as Dragon Dictate) and generating multiple transcripts to represent complex audio and/or video data (using Transana).

Voice Recognition Software

Another particularly useful way to transcribe for those who cannot or should not spend long amounts of time typing is by using voice recognition software. Voice recognition software typically works by being trained to recognize one person's voice, so for transcription purposes the researcher must listen to the recorded data and re-voice it into the voice recognition software package. Transcribing in this way has been discovered by a growing number of researchers (Johnson, 2011; Matheson, 2007), including Paul who uses it for all forms of writing. With Dragon Naturally Speaking (Windows version) or Dragon Dictate (Apple version) voice commands function in the same way that foot pedals and keyboard shortcuts do in other programs.

Visit Web Resource 6.5 to read about Paul's experience of using Dragon.

Using this software will require an investment of both time (to learn the commands and to train the program to recognize your voice) and money (to purchase the software and invest in a quality headset) to get the best results. When it works, however, it is quite impressive. Case Study 6.4 describes how a researcher can use voice recognition software for transcription.

CASE STUDY 6.4

Samuel is a daily user of Dragon Dictate Professional, so it made sense for him to use it for transcription. He mainly does verbatim transcription, using Dragon to type and navigate through the recording. While at his computer, he simply plays the file, and speaks back what he hears. The secret is not to correct as he goes along, but to wait until he has completed three or four paragraphs. Then he uses the playback function, checking what is written on screen with what he hears. Samuel inserts headings into the transcripts using the voice macros. Samuel also used Dragon to make gisted transcriptions by speaking into a Dictaphone. The quality of the headset and the recording device can substantially impact the accuracy of the transcription.

Visit Web Resource 6.6 to learn tips on choosing headsets, eliminating repetitive transcription tasks and setting up macros in Dragon Dictate.

Use of Multiple Transcripts

As audio and video technologies advance, researchers are able to collect ever more complex data sources. Transcription decisions likewise become more complex, as there are additional layers of interaction to represent. For example, to analyse data from a live multi-player computer video game, four separate transcripts would be required to capture 1) the audio interaction, 2) the video feed of the game action, 3) text messages and 4) audio messaging/voiceover, each of

which may be of analytic interest. Software such as Transana, which will be discussed in more detail in Chapter 8, allows multiple video/audio streams and multiple transcripts to be loaded simultaneously for analysis (see Figure 6.13). In addition to capturing the complexity of video game play, these multiple transcripts could be used for analysing focus group data, making it easier to identify speakers and non-verbal behaviour.
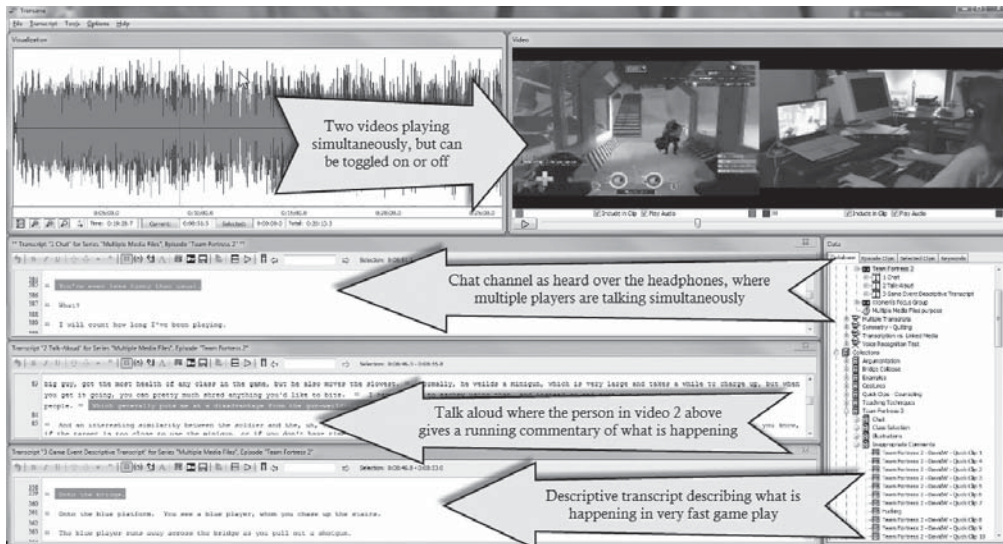

Figure 6.13   Multiple video and transcript features of Transana

Halverson, Bass and Woods (2012) described how they used multiple transcripts to analyse youth films through a variety of analytic lenses: scenes, sound, editing and cinematography.

The first transcript detailed the scenes with, for example, descriptions of where people were standing. The second transcript was a verbatim transcript of everything that was said. The third showed various editing decisions such as camera cuts, zooms and fades. The final transcript highlighted cinematography details such as framing. These transcripts together provided a detailed, nuanced and multi-faceted view into the video file within Transana.

The multiple transcripts in this particular study allowed a depth of analysis that would have been considerably more difficult to achieve without the use of Transana. Their work represents an example of how digital tools can broaden the definition and use of transcripts in qualitative analysis. Take a look at Reflexive Practice 6.6.

——— REFLEXIVE PRACTICE 6.6 ———

Download the free trial version of Transana and use it to practise transcribing your recorded data. Write an entry in your reflexivity journal about your experience testing it out. How might multiple transcripts be useful in your study?

Final Thoughts

The work of the qualitative researcher often includes transforming audio and/or video files into texts through the process of transcription. Each approach to transcription brings with it a certain set of assumptions, and results in a unique representation of the data. Transcription is always a situated act that is selective and representational. The ability to synchronize the transcript with the original media file ensures the transparency and trustworthiness of the research process. While there have been many technological developments to support transcription, there is still no way for a software program to transcribe a media file *for* you unless you are the only speaker. You still have to invest the time in listening and typing or dictating what you hear and choose to represent in the transcript. The development of mobile computing devices has coincided with the development of computer-assisted qualitative data analysis software (CAQDAS) such as NVivo, ATLAS.ti, MAXQDA, Elan and Transana, which will be explored further in Chapters 7 and 8. Many of these programs support direct transcription of media files as well as enabling direct coding of the media files. This has sparked discussions among researchers about whether it is even necessary to transcribe media files prior to analysis (Evers, 2011). While CAQDAS tools certainly have implications for transcribing, they can also have steeper learning curves and be more expensive than the tools described in this chapter which are tailored specifically for transcription.

Chapter Discussion Questions

1. In what ways is transcription a situated act?
2. Compare the four types of transcription – what are the affordances and constraints of each?
3. How might the ability to synchronize the transcript with the audio or video file impact the trustworthiness of the study? How might the ability to represent the recording with multiple transcripts do the same?
4. How can you ensure ethical practices as you transcribe recordings for your study?

Suggestions for Further Reading

To read more about the transcription process, consider beginning with Davidson's (2009) review which reviews over 30 years of literature on transcription practices. Och's (1979) seminal article focuses on the ways in which the transcription process itself is theory driven and is reflective of a researcher's methodological commitments. Poland (1995) is considered a classic text on transcription. To acquire a more thorough understanding of the link between technological development and qualitative methods, including transcription, Lee's (2004) article is a useful beginning point. Evers (2011) is particularly helpful for those who are interested in exploring the possibilities and limitations related to bypassing the transcription process.

References

Ashmore, M. and Reed, D. (2000) 'Innocence and nostalgia in conversation analysis: The dynamic relations of tape and transcript'. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 1(3).

Davidson, C. (2009) 'Transcription: Imperative for qualitative research'. *International Journal of Qualitative Methods*, 8(2).

Dempster, P. and Woods, D. (2011) 'The economic crisis through the eyes of Transana'. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 12(1), Article 16.

Du Bois, J.W. (1991) 'Transcription design principles for spoken discourse research'. *Pragmatics*, 1(1), 71–106.

Evers, J. (2011) 'From the past into the future. How technological development change our ways of data collection, transcription and analysis'. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 12(1), Article 38.

Gumperz, J.J. and Berenz, N. (1993) 'Transcribing Conversational Exchanges', in J.A. Edwards and M.D. Lampert (eds) *Talking Data: Transcription and Coding in Discourse Research*. Hillsdale, NJ: Lawrence Erlbaum and Associates, pp. 91–122.

Hall, T., Lashua, B. and Coffey, A. (2008) 'Sound and the everyday in qualitative research'. *Qualitative Inquiry*, 14(6), 1019–40.

Halverson, E., Bass, M. and Woods, D. (2012) 'The process of creation: A novel methodology for analysing multimodal data'. *The Qualitative Report*, 17(21), 1–27.

Hammersley, M. (2010) 'Reproducing or constructing? Some questions about transcription in social research'. *Qualitative Research*, 10(5), 553–69.

Jefferson, G. (2004) 'Glossary of Transcript Symbols with an Introduction', in C.H. Lemer (ed.) *Conversation Analysis: Studies from the First-Generation*. Philadelphia, PA: John Benjamins, pp. 13–23.

Johnson, B.E. (2011) 'The speed and accuracy of voice recognition software-assisted transcription versus listen-and-type method; A research note'. *Qualitative Research*, 11(1), 91–7.

Kelle, U. (2007) 'Computer-Assisted Qualitative Data Analysis', in C. Seale, G. Gobo, J. Gubrium and D. Silverman (eds) *Qualitative Research Practice*. London, UK: SAGE.

Kvale, S. (2007) *Doing Interviews*. London, UK: SAGE.

Lapadat, J. and Lindsay, A. (1999) 'Transcription in research and practice: From standardization of technique to interpretive positionings'. *Qualitative Inquiry*, 5(1), 64–86.

Lee, R. (2004) 'Recording technologies and the interview in sociology, 1920–2000'. *Sociology*, 38(5), 869–89.

Lester, J. and Paulus, T. (in press) '"That teacher takes everything badly": Discursively reframing non-normative behaviors in therapy sessions'. *International Journal of Qualitative Studies in Education*.

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), 547–60.

Ochs, E. (1979) 'Transcription as Theory', in E. Ochs and B. Schieffelin (eds) *Developmental Pragmatics*. New York, NY: Academic Press, pp. 43–71.

Oliver, D.G., Serovich, J.M. and Mason, T.L. (2005) 'Constraints and opportunities with interview transcription: Towards reflection in qualitative research'. *Social Forces*, 84(2), 1273–89.

Pink, S. (2007) *Doing Visual Ethnography: Images, Media and Representation in Research*. London, UK: SAGE.

Pink, S. (2009) *Doing Sensory Ethnography*. London, UK: SAGE.

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

ten Have, P. (2007) *Doing Conversation Analysis: A Practical Guide* (2nd edition). London, UK: SAGE.