Transcribing Audio and Video Recordings

4

Many researchers dearly wish there was a software that could automatically transcribe recorded interviews. Well, unfortunately, there is still no reliable way of doing this, so in most cases you have to transcribe audio recordings of interviews or focus groups in the traditional way. MAXDQA's Transcription Mode makes this work a great deal easier, but, despite this valuable support, transcribing still takes a lot of time and effort. You will be rewarded, however, by the fact that you can play the original sound at any time and that you can quickly jump to interesting points in the recording via timestamps. As for video recordings, transcriptions may not be absolutely necessary in every case or can be limited to only specific parts of the material.

In This Chapter

- Setting appropriate transcription rules
- Selecting effective transcription settings and starting the transcription process
- Writing partial transcriptions of media files
- Setting and managing timestamps to link transcripts to recordings
- Importing transcripts with timestamps
- Checking completed transcripts

No Transcription Without Rules

In today's research environment, interviews and focus groups are usually recorded with an audio recorder or a smartphone. The recording should be transcribed as soon as possible after the interview so that any potential unintelligible passages can be more easily completed through memories of the spoken words. It depends on the objective of the analysis how detailed and exact the transcription needs to be and

© Springer Nature Switzerland AG 2019

U. Kuckartz, S. Rädiker, *Analyzing Qualitative Data with MAXQDA*,
https://doi.org/10.1007/978-3-030-15671-8_4

which rules should be followed. There are numerous transcription systems (Dresing & Pehl, 2015; Hepburn & Bolden, 2017; Kuckartz, 2014) which differ, among other things, in how particular verbal and nonverbal elements are taken into account in the transcription process. Transcriptions are time-consuming and can therefore also be a cost factor, so it is important to consider what level of accuracy is truly necessary in order to answer your research questions. In most cases, relatively simple transcription systems are sufficient. We defined a set of simple and easy-to-learn transcription rules as part of a previous evaluation project. These software-independent rules are—as supplemented by our own experiences in the meantime and further recommendations by Dresing and Pehl (2015)—listed here:

1. Each speech contribution is transcribed as a separate paragraph. To increase readability, after the paragraph, a blank line is added.
2. Paragraphs for interviewer(s) or moderator(s) are introduced by "I:" or "M:", those for the interviewee(s) by unique abbreviations, e.g., "R:". Numbers are added to the abbreviations ("M1:", "M2:", "R1:", "R2:", etc.) to distinguish between several people in a recording. As an alternative to abbreviations, names or pseudonyms can be used. Labels for speakers are written in bold for better recognition.
3. Speech is transcribed verbatim, i.e., not phonetically or in summary form. Dialects are not transcribed but translated as accurately as possible into the standard form, e.g., standard English.
4. Language and punctuation are standardized slightly where necessary, i.e. to approximate written language. For example, from "He's gonna write a book" to "He is going to write a book." The word order, definite and indefinite articles, etc. are retained even if they contain errors.
5. Clear, longer pauses are indicated by ellipsis points in brackets (. . .). Depending on the length of the pause in seconds, one, two, or three points are used; for longer pauses, a number (in digits) corresponding to the duration in seconds.
6. Intentionally stressed words are underlined.
7. Very loud speech is indicated by writing in capital letters.
8. Affirmative or agreeing utterances made by interviewers (mhm, aha, etc.) are not transcribed so long as they do not interrupt the flow of speech of the interviewee.
9. Short interjections made by the other person, such as "Yes" or "No," are included in brackets in the speech without starting a new paragraph.
10. External interruptions or interferences are noted in double brackets stating the cause, e.g., (cell phone rings).
11. Vocal utterances made by both the interviewee and the interviewer are noted in simple brackets, e.g., (laughs), (groans), or similar.
12. For videos: nonverbal actions are placed in simple brackets, such as (opens the window), (turns away), or similar.
13. Incomprehensible words and sections are identified by (unclear).
14. All information that would allow an interviewee to be identified is to be rendered anonymous.

I: And you're meeting up just now at the end (**R:** Yes, exactly.) or have you been meeting up for a while.

R: Well, during the practice sessions we often had lunch together and did other things but yes, we've only met up for this at the end.

I: Right, right. What grade are you expecting in your exam?

R: You mean as a final grade? (...) I guess probably around a B+ or A-. I mean I'm pretty confident I'll get through it (laughs). So old papers are going around and when I look at them I think, yeah, they're definitely doable. And even just the mock exam took a lot of the anxiety around it out of it for me.

Fig. 4.1 Excerpt from a transcript of an interview

An interview transcribed according to these rules may then look similar to Fig. 4.1. For research papers, the rules used should be documented in the methods section or appendix.

In addition, you might want to consider that you can also transcribe in summary form. In this case, you would not transcribe an interview verbatim but rather summarize the points of interest in longer passages in your own words. This approach can be used if only the content of what is said is of interest and not the wording, and this of course needs to be justified accordingly.

Transcribing Audio Files: Interviews and Focus Groups

To transcribe an audio file, you must first import it into your MAXQDA project, as described in Chap. 3. Once you have imported it, you will see a new text document with a musical note symbol in the "Document System," labeled with the same name as the imported audio file. This is the document in which you will enter your transcription. To start transcribing, right-click on this document in the "Document System" and select ***Transcribe Audio File***. MAXQDA will then switch the display to Transcription Mode, where the audio file is visualized in the "Multimedia Browser" and the transcript in the "Document Browser" directly below it (Fig. 4.2). A transcription settings dialog window will also be displayed each time you launch the Transcription Mode. This window can be closed when it is not required; you can reopen it at any time via an icon in the "Multimedia Browser" when the Transcription Mode is active.

The rewind interval and several transcription-supporting functions can be configured in the settings window. The rewind interval can be set between 0 and 10 seconds, so that this set number of seconds is rewound each time playback is restarted. Usually an interval of 2–3 seconds is a good setting, the optimal value being dependent on the rate of speech, the intelligibility of the recording, and the level of detail required in the transcription. The "Timestamp on enter" setting should always be checked, since this automatically generates a timestamp and inserts a blank line each time you press the enter key in the "Document Browser" (see Rule

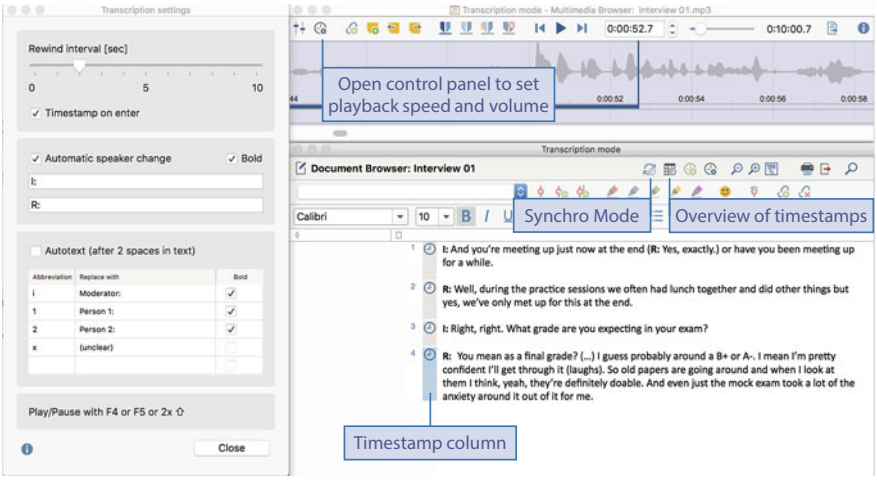


Fig. 4.2 MAXQDA's transcription mode

1 above). The advantage of this is that a section of speech can be played at any time during the analysis by clicking on the corresponding timestamp. Timestamps are used to link the transcript to the audio file such that any part of the original recording can be accessed at any time. Each transcript contains at least one timestamp, set automatically by MAXQDA at the position 0:00:00, which cannot be deleted. Timestamps are displayed in the "Document Browser" as clock icons in a separate column to the left of the transcript (Fig. 4.2). If you hover your cursor over one of these icons, the corresponding time will appear in the tooltip. Clicking on the icon plays the audio file at this point, so that later when you are working with the transcript, you can use the original sound to further interpret and evaluate what was said.

Automatic speaker changes can be switched on when transcribing individual interviews to save yourself typing work. Each time you press the enter key, the next paragraph will start alternately inserting "I:" and "R:", or other freely definable abbreviations for the interviewer and interviewee, optionally in bold (see Rule 2 above). For group discussions, however, you can use the ***Autotext*** option, which means that abbreviations followed by two spaces are automatically replaced by freely definable texts. For each person participating in a group discussion, you can define an appropriate abbreviation, e.g., an i for the moderator, an a for person A, b for person B, and so on. The inclusion of pauses in speech can also be made easier with the autotext option, e.g., by using the 1 for (.) and the 2 for (..).

After all the necessary settings have been made, you can start transcribing. The easiest way to control the audio file and transcription is to use different keys:

- The ***F4*** or ***F5*** keys start playback at the current position—with the rewind interval as set earlier. Press again to pause playback. (The keys ***F4*** and ***F5*** work identically; it doesn't matter which key you use.)

- The same can be done by pressing the ***Ctrl*** key (Windows) or the ***Shift*** ⇧ key (Mac) twice, so that hands can always remain in their original position when typing with ten fingers.
- ***F6*** inserts a new timestamp at the current playback position in the transcript. These manually added timestamps can be used to indicate hard-to-understand, or potentially misinterpretable transcriptions, or other places to be flagged. MAXQDA makes sure that no chronologically erroneous timestamps are added. You can therefore only set another timestamp between two existing timestamps if the current playback position is between the time positions of those two existing timestamps. The current playback area between the last and the next timestamp is highlighted in blue for easier orientation in the timestamp bar.
- ***F12*** *fast-forwards 5 seconds,* ***⇧+F12*** *rewinds 5 seconds, so that you can quickly jump back to unclear sections or find the spot you are looking for more easily.*

The current position in the audio file can always be seen in the "Multimedia Browser." Here you can also jump to a required playback position by using the slider or by directly entering a time.

► **Please Note** On a Mac, function keys ***F1*** through ***F12*** are usually assigned system-specific functions. To use them for transcription in MAXQDA, press and hold the ***fn*** key at the same time. Since this can be very cumbersome, you may want to open the system settings and select the setting ***Use F1, F2, etc. as standard function keys*** in the "Keyboard" menu. Then it will no longer be necessary to press and hold the ***fn*** key. On some Windows computers, the function keys are also assigned to system functions. In this case, it may be helpful to conduct an internet search as to how the function keys on your computer can be used without holding the ***fn*** key.

If it turns out during the transcription that the recorded people are speaking very quickly, you can reduce the playback speed. To do this, click on the ***control panel*** icon in the "Multimedia Browser" (first icon on the left, Fig. 4.2) to open a control window for the volume and playback speed. These settings are also available for playing back when the Transcription Mode is switched off. MAXQDA supports some foot pedals for easier and faster transcription. A list of supported models can be found in the online manual.

To end the transcription, close the "Multimedia Browser" window by clicking on the corresponding icon at the top of the window. MAXQDA then automatically docks the "Document Browser" window with the transcript into its previous position.

► **Tip** To continue with your transcription having closed it, simply right-click on the document and select ***Transcribe Audio File*** again.

Transcribing Video Files

In principle, video files can be transcribed in exactly the same way as described in the previous section for audio files. Right-click on the video document in the "Document System," select the ***Transcribe Video File*** option to activate the Transcription Mode, and adjust your settings for the transcription, if necessary. MAXQDA will display the video images in a window above the waveform visualization of the audio track so you can choose whether to type a description of the images, a verbatim transcription of the speech, or a combination of both.

Occasionally in the case of video files (and sometimes also audio files), only selected parts need to be transcribed, for example, because parts of the recording are irrelevant to the research question or because only particular interaction sequences need to be transcribed and analyzed in detail. For partial transcriptions, you can play the recording by pressing the ***F4*** or ***F5*** key until you reach the point that you want to transcribe; if you already know which point this is, you can of course jump directly to it. Then start the transcription as described above for audio files and continue until the relevant part is over. To keep an overview of the material the whole video covers, it can be helpful to describe the non-transcribed areas, at least in a few keywords, set in brackets in the transcript. If necessary, also type the respective start and end times into the text (Fig. 4.3). Either way, a timestamp should be set at the beginning of each transcribed section, which can be done by pressing the ***F6*** key.

If, while viewing a video in the "Multimedia Browser," you happen to come across a sequence that you would like to transcribe, MAXQDA includes a function for starting transcriptions of selected video clips immediately. To do this, highlight a selection of the video with your cursor in the "Multimedia Browser," right-click on it, and select ***Transcribe*** in the context menu. MAXQA then positions the video at beginning of your selection, switches on the Transcription Mode—if not yet active—jumps to the appropriate position in the existing transcript, and inserts a timestamp there. To help orient you, this timestamp will also be entered as text into your transcript in the form #H:MM:SS-m#. You can then press ***F4*** once, ***F5*** once, or the ***Ctrl*** key (Windows) or the ***Shift*** ⇧ key (Mac) twice in a row to start transcribing immediately.

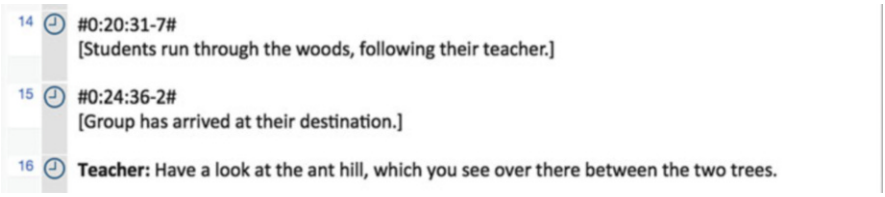
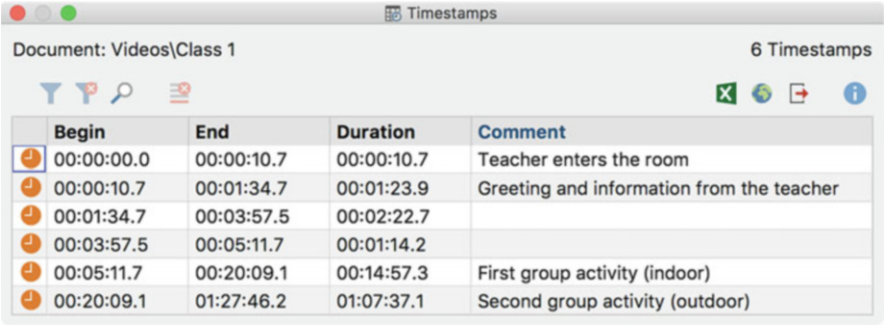


Fig. 4.3 Video transcript with omissions in brackets



Begin	End	Duration	Comment
00:00:00.0	00:00:10.7	00:00:10.7	Teacher enters the room
00:00:10.7	00:01:34.7	00:01:23.9	Greeting and information from the teacher
00:01:34.7	00:03:57.5	00:02:22.7	
00:03:57.5	00:05:11.7	00:01:14.2	
00:05:11.7	00:20:09.1	00:14:57.3	First group activity (indoor)
00:20:09.1	01:27:46.2	01:07:37.1	Second group activity (outdoor)

Fig. 4.4 Overview of Timestamps with comments as an instrument for structuring the transcript

The Overview of Timestamps

As previously mentioned, timestamps are a helpful instrument for linking transcriptions to the original audio or video recordings. MAXQDA also lets you add a comment to each timestamp, in which you can describe the corresponding section of the recording in a few words. In this way, timestamps can help provide a structure to the content of your transcriptions, a tool for partial transcriptions that should not be underestimated. You can add or edit a comment at any time by right-clicking on the timestamp icon. If you hover your cursor over a timestamp, the tooltip that appears will not just show the corresponding time position, but also the comment.

The Overview of Timestamps lists all the times and respective durations of each section (up to the next timestamp) as well as all the comments that have been assigned to these timestamps (Fig. 4.4). You can open the overview by clicking on the icon with the same name in the "Document Browser." This icon appears as soon as an audio or video file is assigned to a text. You can write, edit, and delete comments in the overview, and it is interactive too: double-clicking a row will play the corresponding recorded section, while positioning the text in the "Document Browser" at the same corresponding position in the transcript.

Importing Existing Transcriptions with Timestamps

MAXQDA's transcription functions are perfectly suited for common standard of transcriptions, like the one proposed in Kuckartz (2014, p. 126). Quite often research teams will turn to external service providers to transcribe their data material, who themselves work with specialized programs like Express Scribe, F4, InqScribe, or Transcribe. These programs usually generate transcripts in RTF format, which then need to be imported into MAXQDA together with their corresponding audio or video file. The transcripts usually already contain timestamps entered directly into

Table 4.1 Timestamp formats recognized by MAXQDA

easytranscript, f4 & f5transcript	#hh:mm:ss-x#
HyperTRANSCRIBE	[hh:mm:ss.xxx]
Inqscribe, Transcriva	[hh:mm:ss.xx]
Transana	(h:mm:ss.xx)
Transcribe	[hh:mm:ss]
Transcriber Pro	hh:mm:ss

the text. The timestamps used by the abovementioned transcription tools, for example, are added to the end of each paragraph by default, in the format #HH:MM:SS-m#, where H stands for hour, M for minute, S for second, and m for tenths of a second.

To import a transcript of this kind, select ***Import*** > ***Transcripts with Timestamps***, then select one or more transcripts in the file dialog that appears. If MAXQDA recognizes a timestamp format in the selected document, it will ask you for a corresponding media file to assign to the document and copy to the folder for external files. When you import it into MAXQDA, the existing timestamps will be removed from the document and displayed as timestamp icons next to the text in the "Document Browser." You can also import a transcript with timestamps using all the standard import options in MAXQDA, for example, via ***Import*** > ***Documents***, or by dragging and dropping the document from a folder into your "Document System." MAXQDA automatically checks all text documents to see if they contain timestamps.

MAXQDA automatically recognizes the timestamps used by the transcription programs listed in Table 4.1. If you want to import transcripts from other programs, the timestamp format in these transcripts may need to be adapted. This can be done using, for example, the advanced search and replace functions in word processing programs.

Checking a Transcript

Once you have completed your own transcription or imported an existing transcript, it is usually a good idea to check that the transcription matches the original recording. If necessary, you can increase the playback speed for this process and use MAXQDA's so-called Sync Mode, where the section in the transcript being played is indicated with a color-highlighted background in the "Document Browser." The exact procedure for checking transcripts is explained in Box 4.1.

Box 4.1: Checking a Transcript

- Open the transcript in the "Document Browser," e.g., by double-clicking its name in the "Document System."
- Activate the Sync Mode in the "Document Browser" by clicking on the icon with the same name.
- If it is not already visible, open the "Multimedia Browser" by clicking on the icon with the same name in the "Document Browser." Position the "Multimedia Browser" so that you can see the transcript at the same time.
- Go to the beginning of the audio or video file and start playback.
- If people are not speaking quickly, you can increase the playback speed to hear the transcript faster. To do this, open the ***control panel*** via the icon on the far left of the "Multimedia Browser" toolbar.
- If you find an error in the transcript, pause playback by pressing the ***F4*** or ***F5*** key or by pressing the ***Ctrl*** key (Windows) or the ***Shift*** ⇧ key (Mac) twice. Alternatively, you can skip back 5 seconds to replay the moment in question using the key combination ⇧***+F12***.
- To correct a section of the transcript, switch on the Edit Mode by clicking on the corresponding icon in the "Document Browser," and make the necessary changes to the text. Then switch off the Edit Mode, because the Sync Mode will not work when the Edit Mode is active.

For the Sync Mode to work, several timestamps need to have been assigned to the transcript. If your transcript does not contain any timestamps, except for the very first (MAXQDA generates this one automatically, starting at 0:00:00), there is no need to activate the Sync Mode. In this case simply make sure that you set the playback speed at a reasonable rate.

References

Dresing, T., & Pehl, T. (2015). *Manual (on) transcription: Transcription conventions, software guides and practical hints for qualitative researchers* (3rd ed.). Marburg: Dr. Dresing & Pehl GmbH.

Hepburn, A., & Bolden, G. (2017). *Transcribing for social research*. Thousand Oaks, CA: SAGE.

Kuckartz, U. (2014). *Qualitative text analysis: A guide to methods, practice & using software*. Los Angeles: SAGE.