Qualitative Data Transcription and Translation
IRC Research Toolkit

This document aims to provide guidance to IRC staff on transcribing and translating recordings from interviews for research, monitoring and evaluation.

Transcription

Transcription is the action of providing a written account of spoken words. In qualitative research, transcription is conducted of individual or group interviews and generally written verbatim (exactly word-for-word).

Transcribing may appear to be a straightforward technical task. However, the process of transcription may differ depending on its end use[1]. Transcripts that are used mainly to select quotes and sound bites may not need the same level of details as transcripts which will be systematically reviewed, grouped into themes (often through a process of coding), and analyzed for content.

The sections below provide guidance on 1) whether to transcribe, 2) budgeting time and resources required for transcription, 3) hiring transcribers, 4) tips and best practices in transcription, and 5) ethics and confidentiality.

1) Deciding Whether to Transcribe

Is it always necessary to transcribe audio or audio-visual recordings? Not always. Some qualitative data analysis software allows users to code sound recordings (as opposed to written text). However, keep in mind that this won't work if the recordings require translation. Also, some researchers prefer to have written transcripts when conducting analysis, either for ease of use or to have a back-up in case of technical failure.

2) Budgeting Time and Resources

Transcription is a detailed, time-consuming activity, especially when done well. It is easy to under-budget time and resources. It can take anywhere from three to 10 hours to transcribe one hour of recording, depending on several factors including the level of detail required, the transcriber's skill, and the quality and complexity of the recording (e.g., whether several participants are talking at once). For guidance on budgeting sufficient resources and time, you can consult the Qualitative Methods Budget Template.

Helpful Hint

As a general rule of thumb, focus group discussions take longer to transcribe than interviews with a single participant.

Timing of transcription is also important. Transcribing should occur soon after an interview / focus group discussion, when the discussions are still salient. This way, if recordings are unintelligible etc., transcribers may consult with the interviewer to clarify what was said.

[1] Mondada, Lorenza. 2007. "Commentary: transcript variations and the indexicality of transcribing practices".

If resources allow, it is considered best practice to first transcribe an interview in the original source language and then translate it into the target language. This way you can continually check transcripts against translated interpretations during analysis.

Tips on Recording Interviews

- Purchase recording devices with built-in USBs to facilitate easy transfer of data from the recording device to a computer.
- Test the recording devices before starting data collection. Check for sound clarity and compatibility with your computers.
- If possible, always equip interviewers with extra batteries, an extra recording device, paper and pens/pencils as a back-up.
- Always write notes (even if mostly bullet points), even when recording device is available. If a recording fails or is not sufficiently clear, at least you'll have some data to analyze.
- Ensure the data collection activities are carried out in a quiet location (for recording purposes).
- Immediately after qualitative data collection, spot check the quality of the recording. If the quality is poor (e.g., voices are muffled), immediately jot down as much detail about the discussions as possible.
- In some cases, it may not be appropriate to record the interview (e.g., when discussing particularly sensitive topics, when capturing nuanced detail is unnecessary).

3) Hiring Transcribers

- When hiring a transcriber, researchers should consider a number of factors, including the person's native language, computer skills and typing speed, attention to detail, as well as his/her familiarity with the population being studied and the vernacular being transcribed.

- Candidates who are being considered for a transcription position should be given a test – i.e. to listen and transcribe a mock recording developed by the researcher in an allotted timeframe – to assess his/her transcription abilities.

- Prior to this assessment, candidates should be provided with specific instructions about the research purposes, and the corresponding transcription requirements (e.g., expected structure and level of detail).

- Especially for new staff, perform spot checks of transcriptions while the process is ongoing. Listen to portions of the recording and check this against the transcript for accuracy. If the recording is in a language which you do not understand, ask a trusted, experienced staff member who does.

- Another technique is to have multiple people transcribe the same material and to compare the different versions.

4) Tips and Best Practices in Transcription

Decisions about the level of detail (e.g., whether to transcribe or omit non-verbal dimensions, like body language, or involuntary vocalizations), and data representation (e.g., representing the verbalization "hwarryuhh" as "How are you?") should be discussed with transcribers in advance. There is an important trade-off to be made between the accuracy of the transcript and its readability. How much detail is enough? This depends on the research aims. Here are a few tips:

- Unintelligible text or silence should be noted, with the approximate seconds listed. Similarly, if transcribers are unclear about the precise wording (e.g., if speech is slightly muffled), this should be flagged (e.g., with red text) and the transcriber should follow up with the interviewer for clarification.

- Particularly for case studies, “thick” descriptions are very important. You may want to consider transcribing visual information that will help with data interpretation, for example, room layout, body orientation, facial expression, gesture and the use of equipment in consultation. In other types of analyses, this level of detail may be too time-consuming for its worth.

- Decide whether full representation is important. Certain representations of accents may be unhelpful, either because it can make the text difficult to read, because it may make the speaker sound less educated, or convey their ethnic background or place of origin.

- Decide whether it’s important to transcribe certain sounds like “Hm”, “Ok”, “Ah”, “Yeah”, “Um”, “Uh”, and “Uh huh” / “Nuh uh”. Unlike many involuntary noises, like throat clearing, there is meaning attached to them that can influence a conversation. Often, these vocalizations are not transcribed, but can provide a great deal of insight into both the nature of conversation (i.e., how one converses), but also the informational content of the conversation (Gardner 2001).

- Once you have made decisions on the above, it is good practice to provide transcribers with a specific notation system and examples to use as a guide.

5) Ethics and Confidentiality

- Before starting data collection, check whether you are required to obtain ethical approval from an Institutional Review Board. This is required if you are conducting research with humans.

- If you are hiring a transcriber, it is recommended that he/she sign a confidentiality agreement to prevent the disclosure of participants’ personal information.

- Researchers will need to determine when personally identifying information (PII) should be removed from the transcript. This may be at the time of transcription, in which case, transcribers should be informed of what constitutes PII and what placeholders should be used (for instance, with the following notation: [city street]). If in doubt, consult the overseeing Institutional Review Board for advice.

- As explained under the “Tips and Best Practices in Transcription” above, highly accurate transcripts (with representations of accents and verbal idioms) may risk exposing participants’ identities. Consider this risk when instructing transcribers on the level of detail and representation required for the purposes of the analysis.

Translation

Transcription that involves translation from one language to another presents an especially complex and challenging task. If the researcher is not a native speaker of the language used by research participants, translators are often required. This section provides guidance on 1) budgeting time and resources required for transcription, 2) hiring transcribers, 3) tips and best practices in transcription, and 4) ethics and confidentiality.

1) Budgeting Time and Resources

Depending on skills of transcribers and the relevant languages, time & effort to transcribe can vary quite broadly, from an estimated 3-10 hours to transcribe each hour of transcript. Calculate an estimated 1 hour of coding per person for every page of transcript. For translation, a rule of thumb is to estimate that 10 pages can be translated per day per person.

2) Hiring Translators

- The quality of data translation can have significant implications on the conceptual equivalence and

accuracy of study findings. Therefore, choosing and evaluating translators is critical. The translator not only literally translates language, but also adds cultural or interpretive insights.

- Researchers may choose to hire a certified translator. However, in the contexts in which we work, this is often infeasible. It is important to take into consideration candidates' experience, education level, whether he/she is a bilingual native speaker, his/her socio-linguistic competence, and whether he/she is from the same area as subjects.

- Candidates who are being considered for a translator position may be given a test – i.e. to listen and translate a mock recording developed by the researcher in an allotted timeframe, then have it back-translated by another native speaker.

- In order to adequately budget the translators' time, investigators should systematically plan for how much to involve translators in the research process (e.g., in question development, data collection, data analysis, and dissemination).

- Especially if this translator is a new staff person, it could be important to spot check the quality of the translation as it is ongoing by asking a native speaker to provide an alternate translation or a back-translation.

- Brislin's model of translation is considered by many to be best practice and involves recruiting at least two bilingual people who translate the qualitative research recordings from the original source language into the target language. Another bilingual person will then back-translate the documents from the target language to the source language, and finally both versions are compared to check accuracy and equivalence. Any discrepancies that have occurred during the process are then negotiated between the two bilingual translators.

- The objectivity of the interpreter should be stressed in trainings (e.g., not modifying answers based on what the translator thinks the researcher wants to hear).

3) Tips and Best Practices in Translation

- Research questions should be developed in the source language without the use of slang, colloquialisms, and complex sentence structures to avoid errors when translated.

- Translated questions should also be pilot tested before undertaking the main qualitative study.

- Researchers may find it useful to develop a translation lexicon to serve as a consistent resource for the translator. For example, codebooks for analyses could be provided to the translator in both the target and source language. This could include tips on:

    - What to do when words do not have an exact equivalent translation. For example, in this case the original word may be maintained in quotes with the closest description in brackets or footnotes. It may be a good idea to discuss the meanings of particular words amongst the study team.

4) Ethics and Confidentiality

- Before starting data collection, check whether you are required to obtain ethical approval from an Institutional Review Board. This is required if you are conducting research with humans.

- If you are hiring a translator, it is recommended that he/she sign a confidentiality agreement to prevent the disclosure of participants' personal information.

Additional Resources

The above guidelines are drawn from the author's experiences and the following (open source) recommended readings:

- Bailey J. (2008). First steps in qualitative data analysis: transcribing. *Family Practice*, 25, 127–131.
- Davidson, C. (2009). Transcription: Imperatives for Qualitative Research. *International Journal of Qualitative Methods 2009, 8(2).*
- Oliver, D., Serovich, J., & Mason, T. (2005). Constraints and Opportunities with Interview Transcription: Towards Reflection in Qualitative Research. *Soc Forces*, 84(2), 1273-1289.
- Regmi, K., Naidoo, J., & Pilkington, P. (2010). Understanding the Processes of Translation and Transliteration in Qualitative Research. *International Journal of Qualitative Methods 2010, 9(1).*
- Squires, A. (2008). Language barriers and qualitative nursing research: methodological considerations. *Int Nurs Rev*, 55(3): 265–273. doi:10.1111/j.1466-7657.2008.00652.x.
- Temple, B. & Young, A. (2004). Qualitative research and translation dilemmas. *Qualitative Research, 4(2)*, 161-178.