

# social research Update

## Using NVivo Audio-Coding: Practical, Sensorial and Epistemological Considerations

### Megan Wainwright and Andrew Russell

Megan Wainwright, a PhD student in Anthropology, is the Van Mildert College Trust Research Scholar at Durham University, undertaking ethnographic research on Chronic Obstructive Pulmonary Disease (COPD) in Uruguay. She has researched and published in the areas of medical education, palliative care, end-stage renal disease and COPD at McGill University. The research on chronic constipation discussed here won the RAI Arthur Maurice Hocart Prize 2009 and the Dryburn Research Prize 2010. <megan@meganwainwright.ca>

Dr Andrew Russell, a Senior Lecturer in the Department of Anthropology at Durham University, is Director of the University's Community Outreach and Engagement programme. <a.j.russell@durham.ac.uk>

- Social researchers have depended on verbatim transcripts of audio-recorded interviews in order to apply rigorous and systematic analyses. With software programs such as NVivo8, transcription is no longer a necessity.
- This paper describes NVivo8 audio-coding for open-ended interviews, based on the narrative-ethnographic study of an outpatient clinic.
- While thematic narrative analysis was the approach applied to the interviews, the process described and issues raised are relevant to any form of analysis undertaken using programs similar to NVivo8.
- The use of this technology raises questions about the epistemological and sensorial differences between visual and audio interpretation and analysis.

The semi-structured or unstructured open-ended interview is a seminal method in qualitative social research. In many research fields, the interview offers access to individuals' own interpretations, perceptions, experiences and practices. Transcripts, the verbatim write-up of audio-recorded interviews, are a respected and in many ways fetishised form of data, in part because they lend themselves to systematic content-coding and analysis.

However, voluminous transcripts are produced as something of a 'knee-jerk' response to data creation, with most of the material they contain left to gather dust in the researcher's drawer. Furthermore, transcription

is a laborious process and it is increasingly common that transcriber and analyst (and sometimes even the interviewer) are different people (Mishler 1984). Producing a transcript can be seen as the first step in the analytical process, assuming that the researcher and transcriber are the same person. When they are not, at the very least the researcher should check transcripts against audio recordings to correct any transcription errors, which is a lengthy process in itself.

Here we demonstrate how one can work directly with audio data throughout the analytical process; i.e. without using a written transcription. Above all, this *Update* calls for

critical reflection on the centrality of transcripts and the relationship between mode of engagement with data and analytical thought about them.

### The Project

NVivo8 is a qualitative data-analysis software program released in 2008 that enables the coding and analysis of text, image, audio and video data. The specific features of NVivo8 and how to use them have been discussed elsewhere (Auld et al 2007; Bringer et al 2006). However, the ability to directly code audio data is a new feature of NVivo8 and has not yet been discussed in the literature. While programs such as ATLAS.Ti (Hwang 2008), Transana (specific to video and audio) and MAXQDA also support audio data, this *Update* focuses only on NVivo.

One aim of the MSc research on which this paper is based was to test whether the analytical process of thematic coding, previously used in the analysis of text-based data, could be replicated by working exclusively from audio-recordings using the applications available in NVivo8. Eleven unstructured narrative interviews (lasting between one and three hours) exploring the illness experience and healthcare relationships of people receiving specialist care for chronic constipation were used. The theoretical orientation of the project was that of grounded theory, and analysis was informed by two similar approaches to narrative-analysis: the holistic-content perspective (Lieblich, Tuval-Mashiach and Zilber 1998) and thematic narrative analysis (Riessman 2008). As discussed elsewhere, qualitative data analysis software programs lend themselves to grounded theory (Bringer et al., 2006; Peters and Wester 2007), although MacMillian and Koenig (2004) offer a thought-provoking critique of this association.

Exclusive use of audio-coding offered significant advantages. Firstly, the

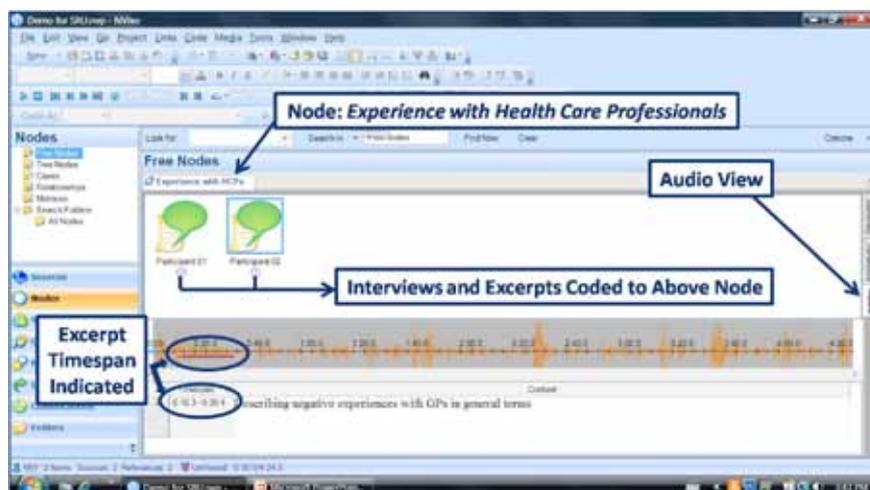
fieldwork was of five months' duration and it would have been unfeasible for the researcher to transcribe all the interviews in that time-frame in addition to analysis and writing-up. Because of the participatory nature of the research it was considered important that each interview be given equal analytical attention (i.e. we did not want to select only a sub-sample for transcription and analysis). Secondly, the narrative interviews elicited very intimate details of participants' lives and it was felt inappropriate to outsource transcription to a third party.

### Using Audio Coding – an Example

Audio-coding in NVivo8, like document (text) or video coding, requires the creation of 'Internal Source' files within the 'Sources' tab. For example, in this study eleven audio 'Internal Source' files were created by importing the interview recordings (WMA format) into NVivo8 and labelling them according to participant code. Once the 'Internal Sources' have been created, a number of applications become available such as Nodes, Memos, Sets, Queries, Models, Links and Classifications. Bazeley (2007) gives a detailed description of these and other features of the NVivo program. Nodes are essentially codes – the

topics or themes created by the user to which data-chunks (i.e. parts of an audio recording or sections of a written document) can be coded. An imported audio-file will appear as an Audio Wave Form which can be listened to and divided up into audio-excerpts.

An audio-excerpt is assigned a 'Timespan', notes can be typed into the 'Content' column (to help identify the excerpt) and then is assigned to a Node. For example, 1.55 min to 2.30 min (Timespan) was coded as 'Family Relationships' (Node) with the note "discussing the effect the illness has on marriage" (Content). The creation of Nodes is dynamic: as more interviews are coded, Nodes can be added, changed, collapsed or deleted. The second step is to create a 'Memo' - a short summary and initial analytical reflections linked to that particular 'Internal Source'. One can re-listen to the excerpts coded under each Node for emerging themes. The way one navigates the various Nodes is the same as for textual data and is explained in Bazeley (2007). However, one noteworthy feature is that from within each Node a hyperlink opens the corresponding interview and highlights the relevant sections along the Audio Wave Form, making it easy to return and listen to audio-excerpts.



## Practical, Sensorial and Epistemological Considerations

Audio features offer time and budgetary savings; advantages software companies certainly advertise. We believe that using the audio features to engage in audio-analysis is also attractive for analytical reasons and raises interesting questions concerning the relationship between a researcher and her data that need to be explored further. The software facilitates working with audio according to the same processes we are accustomed to with transcripts: systematic and rigorous analysis, enabling one to work through an entire interview without neglecting sections, annotating and memo-ing along the way etc. It also has the advantage of enabling the researcher to move swiftly between codes and audio excerpts, keeping the researcher close to the original data. However, a number of challenges and considerations arise which will be addressed here. Some of these challenges are practical ones involving how best to use the program and getting over our habituation with texts. Deeper challenges arise because of epistemological and sensorial differences between thinking critically and analytically about what we read compared to what we hear. These important questions can only be refined and explored if and when more social researchers start audio-coding some of their interviews and comparing their experiences.

## Practical Challenges

Compared to transcripts, audio data may feel less tangible and it can be harder to review and locate particular utterances (e.g. for inclusion in an article). However, this mostly has to do with the effectiveness of the annotations included in the 'Content' column for each audio-excerpt. Annotating an audio-excerpt appropriately and avoiding the trap of half-transcribing everything one hears, is a skill that

requires a good deal of practice. One could argue that annotating is a form of transcription, but in our experience the most difficult process is to annotate briefly so that the researcher is drawn back to listening to the excerpts as opposed to reading a summary of what was said. For instance, if one is coding thematically one can type in the content column what is happening or what the participant is "doing" in that period of time in the interview rather than writing out everything that was said (e.g. "participant describing relationship with general practitioners"). By then coding all such passages across interviews as "healthcare relationships" the researcher then has to re-listen to each excerpt under that Node to begin to analyze across interviews. Admittedly we often fell into selective transcription. However, this was often a result of knowing that the research output would be text (a dissertation) and therefore wanting to type out seminal quotes and facilitate being able to find them. In retrospect, what we needed were better ways of flagging important utterances either by creating memos with seminal quotes or creating a node called "seminal quotes". These issues need further investigation particularly with regards to various analytical approaches (e.g. thematic vs. discourse analysis) and types of data (e.g. interviews vs. focus groups).

## Epistemological and Sensorial Challenges

In challenging the necessity of transcripts for the systematic analysis of interviews, we begin to think critically about the way in which interviews only gain their legitimacy as 'data' once they are transformed into text. There is a gap between speech and text whereby written versions of speech cannot capture all that spoken words communicate (Mishler 1984). With audio-coding, this gap is narrowed during the

analytical process as the researcher is "sensorially" (Nichter 2008) closer to the data. The researcher can think analytically about the data while being immersed in the flow of the recorded interview, attending to utterances, silences, emotions and the interactive dialectic between interviewer and interviewee in ways that are difficult when reading even detailed transcriptions. Of course, an interviewer returning to an interview transcript may find it elicits sensorial memory. The phenomenon of sub-vocalization – when we hear in our minds the words we read – might play a part in bringing us back to that interview, that house, that day etc., assuming we were there in the first place. However, this may not be as effective as hearing the recorded voices all over again while engaged in analytical reflection rather than transcription or transcript checking. Another option, possible with NVivo8 and other programs, is to synchronize audio and transcript so that one can read the transcript while audio-coding or vice-versa. Whether this time-consuming process is analytically superior to just working from text or audio alone is also something that needs to be explored.

Another point is that with audio-based analysis the researcher may feel less confident about where and when to delineate an audio-excerpt. With a transcript, the analyst can generally draw boundaries wherever she sees fit. When coding interviews auditorily, we found boundaries seemed to coincide more naturally with the starts and stops of the conversation. It is also more difficult to stop and rewind an interview than it is for our eyes to move back and forth across a printed page. What would facilitate this is a transcription foot pedal to control movement along the Audio Wave Form.

One could argue that the written transcript allows for a certain detachment which in the analytical process can be helpful. On the other

hand, analyzing an interview as spoken conversation as opposed to text may have analytical benefits in terms of remaining in-context and respecting coherence and narrative flow which in turn might strengthen interpretation and analysis.

Within cognitive psychology there is still much to learn about how our brains work when analysing visual/written words compared to heard speech but undoubtedly the neurological pathways and processes involved are different. Thinking analytically about an interview and coding it into parts while listening is significantly different from doing it by reading. McLuhan argues that "Western man thinks with only one part of his brain and starves the rest of it. By neglecting ear culture, which is too diffuse for the categorical hierarchies of the left side of the brain, he has locked himself into a position where only linear conceptualization is acceptable" (2004:69). Similarly, Stoler says sight is "the privileged sense of the West" (1989:5), causing the neglect not only of hearing but all other senses. According to Classen, in his historical and cross-cultural work on the social construction of sensory perception, Western sight is the "sense of science" (1993:6). These arguments may account for the likely resistance to audio-coding, and our unquestioned dependence on texts. They also suggest that some degree of re-training may be necessary to overcome our shortcomings and feel confident working in aural modalities.

## Conclusion

NVivo8 and similar programs offer social researchers opportunities to engage analytically with their data in new and exciting ways and the opportunity to reconsider and challenge the taken-for-granted necessity of transcription. Doing so enables further consideration of how

the mode of engagement with data may affect the ways in which it is assimilated and analysed. However, despite the potential for narrowing the gap between speech and text that audio-coding permits, the final product, usually a written paper, falls short in this regard. The next chapter in the development of audio technologies for research is surely to push researchers to think more about the audio-diffusion of results and research findings via pod-casts, radio segments, interactive online publications and other innovative methods that challenge the visual/textual dominated *status quo*.

## References

- Auld, G. W., et al. (2007), 'Development of a decision tree to determine appropriateness of NVivo in analyzing qualitative data sets', *Journal of Nutrition Education and Behavior*, 39 (1), 37-47.
- Bazeley, Pat (2007), *Qualitative data analysis with NVivo* Sage Publications.
- Bringer, J. D., Johnston, L. H., and Brackenridge, C. H. (2006), 'Using computer-assisted qualitative data analysis software to develop a grounded theory project', *Field Methods*, 18 (3), 245-66.
- Classen, Constance (1993), *Worlds of sense: Exploring the senses in history and across cultures* Routledge.
- Hwang, S. (2008), 'Utilizing Qualitative Data Analysis Software A Review of Atlas.ti', *Social Science Computer Review*, 26 (4), 519-27.
- Lieblich, Amia, Tuval-Mashiach, Rivka, and Zilber, Tamar (1998), *Narrative research: Reading, analysis, and interpretation* (Applied Social Research Methods Series, 47) Sage Publications, 187.
- MacMillan, K. and Koenig, T. (2004), 'The wow factor - Preconceptions and expectations for data analysis software in qualitative research', *Social Science Computer Review*, 22 (2), 179-86.
- McLuhan, Marshall (2004), 'Visual and acoustic space', in Christopher Cox and Daniel Warner (ed.), *Audio culture: Readings in modern music* The Continuum International Publishing Group, 67-72.
- Mishler, Elliot G. (1984), 'Approaches to research on medical interviews: A critical review', *The discourse of medicine: Dialectics of medical interviews* Ablex Publishing Corporation, 17-58.
- Nichter, Mark (2008), 'Coming to our senses: Appreciating the sensorial', *Transcultural Psychiatry*, 45 (2), 163-97.
- Peters, V. and Wester, F. (2007), 'How qualitative data analysis software may support the qualitative analysis process', *Quality & Quantity*, 41 (5), 635-59.
- Riessman, Catherine Kohler (2008), *Narrative methods for the human sciences* Sage Publications.
- Stoller, Paul (1989), *The taste of ethnographic things: The senses in anthropology* University of Pennsylvania Press.

---

*social research UPDATE* is distributed without charge on request to social researchers in the United Kingdom by the Department of Sociology at the University of Surrey as part of its commitment to supporting social research training and development.

Contributions to *social research UPDATE* that review current issues in social research and methodology in about 2,500 words are welcome. All *UPDATE* articles are peer-reviewed.

