



social research UPDATE

- The technology needed to make digital recordings of interviews and meetings for the purpose of qualitative research is described.
- The advantages of using digital audio technology are outlined.
- The technical background needed to make an informed choice of technology is summarised.
- The Update concludes with brief evaluations of the types of audio recorder currently available.

Tools for Digital Audio Recording in Qualitative Research

Alan Stockdale

In a recent book Michael Patton writes, “As a good hammer is essential to fine carpentry, a good tape recorder is indispensable to fine fieldwork” (Patton 2002: 380). He goes on to cite an example of transcribers at one university who estimated that 20% of the tapes given to them “were so badly recorded as to be impossible to transcribe accurately – or at all.” Surprisingly there is remarkably little discussion of tools and techniques for recording interviews in the qualitative research literature (but see, for example, Modaff and Modaff 2000).

This overview discusses the potential advantages of digital recording and provides some technical background and a checklist of features to consider when buying a digital recorder. It concludes with brief comments on the different types of recorder currently available and the names of some of the leading manufacturers. As the technology is changing rapidly and new recorders are appearing constantly there is little point in recommending particular models as the recommendations would rapidly become obsolete.

Detailed discussion of the methodological issues associated with audio recording is beyond the scope of this *Update*. Let it be noted, however, that audio recording uses technology and technique to frame and structure the representation of an event. It is important to keep this in mind given that the quality improvement of digital technology, when coupled with technical

naïveté, can heighten the sense of “being there”. For discussion of the naturalization of audio recordings in qualitative research, see Ashmore and Reed (2000).

Why digital?

Audio Quality

The recording process used to make analogue recordings using cassette tape introduces noise, particularly tape hiss. Noise can drown out softly spoken words and makes transcription of normal speech difficult and tiring. Digital recorders generally have a much higher signal to noise ratio. Less noise reduces the risk of lost data and results in faster, less expensive and more accurate transcription.

Note that audio quality also depends on using a suitable external microphone or microphones properly positioned near speakers in an environment with low levels of ambient noise.

Digital Editing

There are cheap, sophisticated audio editing programs (e.g. Syntrillium’s CoolEdit 2000) that can be helpful if they are used with care. These programs can be used to adjust the recording level, fix recordings in which one speaker sounds louder than another, reduce unwanted background noise, filter unnecessary frequencies, silence personal or identifying information to protect anonymity, and cut extraneous sections from the beginning or end of audio files.

Dr. Stockdale’s training is in cultural anthropology. He is a senior research associate at Education Development Center in Boston, Massachusetts, where he currently serves as an investigator on several genetics education research projects funded by the U.S. National Institutes of Health.

social research **UPDATE**

Archiving

It is easy and inexpensive to backup and archive digital audio recordings. When using a compressed digital format such as MP3 it is possible to store an entire research project on one or two CD-ROMs. However, because digital audio is readily copied and transmitted, additional steps may need to be taken to ensure that original recordings are kept secure and research participants' confidentiality is adequately protected.

Computer-based Transcription

Transcription software can greatly improve the usability and usefulness of transcriptions (Muhr 2000). This is primarily accomplished through the automatic insertion of tags that encode additional data during transcription. See for example DGA's Transcriber software, which uses eXtended Markup Language (XML) tags. The most obvious use of tags is synchronization of transcript files to their corresponding audio files. This facilitates checking, correcting, or later referral as the researcher can go to any point in the transcript and immediately play the corresponding segment of audio. Hopefully, these type of features will be integrated into analysis software eventually.

Technical background

Frequency Response

The audible range of the human ear is approximately 20 Hz to 20 kHz. The most important frequencies for speech occur in the "mid range" between 250 Hz and 8 kHz. The sensitivity of microphones and recorders to audio frequencies varies. Microphones for recording speech often are most sensitive to the range between 200 Hz and 10 kHz. Digital recorders also vary in their sensitivity. A MiniDisc recorder, when matched with an appropriate microphone, is capable of recording frequencies between 20 Hz and 20 kHz. Some digital voice recorders when set to "long play" mode may only encode frequencies between 300 Hz and 3 kHz. Telephone line frequencies are limited to those between 400 Hz and 3.4 kHz. A frequency response that approximates the mid-range frequencies will result in the best speech recordings.

Channels

Single channel or mono recording often works fine for interviews. Mono recording also doubles the available record time when using a digital recorder. However, stereo recording may be an advantage in some situations where the speakers are separated from each other or where there are several speakers. To take advantage of stereo recording a microphone setup that allows each microphone element to be positioned next to a different speaker or set of speakers will be necessary. This will aid transcription by making it easier to ensure that a good audio recording level is obtained for all speakers and making it possible for the person doing the transcription to use the stereo separation to help identify speakers and transcribe overlapping speech.

Recording Level

The level of the audio signal – how much the microphone signal is amplified – needs to be set properly to make a good recording. If the signal is too strong it will be distorted; too weak and the speech one wishes to record may be swamped by noise and difficult to hear. The majority of cheap recording devices do not provide any visual display of the level and set the recording level automatically. This makes recording easy but automatic level control (ALC) can be problematic (Modaff and Modaff 2000). ALC constantly adjusts the level to any audio input, even background noise during pauses in speech. This may result in the level being frequently, although briefly, poorly adjusted to the speech being recorded. ALC also changes the overall dynamics so that the difference between loud and quiet speech is compressed.

Digital Audio

Digital audio is recorded by sampling a sound wave and assigning each sample a value. The quality of the audio depends on the sampling frequency and the resolution, that is, the range of values that can be assigned to each sample. The sampling frequency is significant to the extent it needs to be at least double the highest frequency one wishes to record. Music CDs use a sample rate of 44.1 kHz – a rate more

social research UPDATE

than adequate for encoding frequencies up to 20 kHz. For recording speech, a sample rate of at least 16 kHz will ensure good quality. Audio is normally encoded in 8 bits, 16 bits, or in some cases higher resolutions. The higher the bit depth, the greater the number of amplitude values that can be represented for each sample. An 8 bit resolution may be adequate for recording speech for some purposes, but 16 bits is better.

Digital Audio and Compression

CD quality digital audio corresponds to a sample rate of 44.1 kHz, encoded at 16 bits, on two channels. This works out to:

$44.1\text{k} \times 16 \times 2 =$	1411.2	kilobits/second
$1411.2\text{k} / 8 =$	176.4	kilobytes/second
$176.4\text{k} \times 3.6\text{k} =$	635.04	megabytes/hour

To record at this rate consumes a considerable amount of storage space. The same is true of other forms of Pulse Code Modulation (PCM) audio, which is the usual format for Windows WAV files and Macintosh AIFF files. Even if the encoding rate is reduced by using a 16 kHz sample rate at eight bit resolution with one channel (which for many purposes might be satisfactory for recording speech) the recording will still consume 57.6 MB/hr.

The solution to the space problem is compression schemes or codecs that use psychoacoustic principles and other audio features to reduce the bit rate in ways that limit the perceived quality loss of the audio stream. Common compression schemes include: Fraunhofer MPEG 1 Layer 3 (MP3), Advanced Audio Coding (AAC), Adaptive Transform Acoustic Coding (ATRAC), and Windows Media Audio (WMA). MiniDisc uses ATRAC, which in standard mode, like CD audio, samples at 44.1kHz, in stereo, and encodes in 16 bits, but saves the audio in 1/5 the space without perceptible loss of quality. Fraunhofer MP3 saves audio in 1/11 the space. A Fraunhofer MP3 audio file encoded at 32 kbps (22.05 kHz sample rate, with 16 bit encoding, mono) will provide good voice recording for many purposes and only takes up 14.4 MB/hr. Newer codecs such as WMA and AAC maintain perceived audio quality at even greater compression ratios.

Factors to consider

- Cost (including batteries and media if applicable). It is a false economy to purchase a cheap recorder if the audio quality is such that it increases the cost and time of transcription. Transcription is time consuming and expensive so a good recorder costing hundreds of pounds will quickly pay for itself.
- Audio quality. See discussion above.
- Easy of use.
- Microphones. Are external microphones supported and are they supported through line-in or mic-in jacks? Internal microphones are usually of low quality, may pick up noise from the recorder, and may be difficult to position close to interviewees. An external amplifier, often expensive and cumbersome, may be needed to boost the microphone signal to line level if an external microphone can only be attached through a line-in jack.
- Portability and intrusiveness in an interview situation.
- Ruggedness and reliability of recorder and media.
- Audio formats. What type of recording formats or compression schemes are supported? How easy is it to use a given format with data analysis, audio editing, and transcription tools?
- Stereo recording. Is stereo recording supported? See discussion above.
- Computer transfer. How easy is it to transfer recordings to a computer? Is USB or some other method that allows faster than real time upload supported?
- Record time. How long will media and batteries allow recording to continue uninterrupted?
- Batteries. Can rechargeable batteries be used?
- Control over the recording process. Is the recording level displayed and is it possible to manually adjust the recording level? See discussion above.
- Additional information display. Does the recorder display remaining battery power and remaining record time?
- Copy protection. Is a copy protection scheme implemented and, if so, what

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.

social research UPDATE

limitations does it impose on the use of audio recordings?

Types of digital recorder

Many of the digital recorders designed specifically for recording interviews or meetings are expensive, complicated, and geared to the needs of broadcast journalists. Other types of digital recorder that are simpler and cheaper are often designed primarily as portable music players or for simple dictation and may have some significant limitations when used to record interviews and meetings. At the moment, there are few devices that fall in the middle ground, but new ones are constantly appearing.

Cassette Tape Recorder (Sony, Marantz)

While not digital, a cassette recorder can still be used to create digital audio files by re-recording cassette tapes to a computer equipped with a soundcard. Disadvantages are a low signal-to-noise ratio, limited recording time, and the need for analogue to digital conversion.

Handheld Computers (Handspring, Palm, HP, Toshiba, NEC, Casio)

PocketPC and Palm devices can be used to record audio but very few of these devices support the use of external microphones. Handheld computer devices may eventually appear with input jacks or add-ons that allow external microphones to be used.

Desktop and Portable Computers

Direct to computer recording may be the best and cheapest way to make digital recordings of interviews done by phone when equipped with a good telephone coupler, soundcard or USB audio input device (e.g. Griffin Technology's iMic), and recording software. A computer may be cumbersome for field recordings. The latest ultra subnotebooks (Sony, Toshiba, Fujitsu) are quite small and light, but availability is often limited outside Japan. This is an expensive option but most people either own a computer already or need one for other tasks.

Consumer Player/Recorders

Some portable consumer devices that are primarily designed for listening to music can be used to record speech. At the moment, these devices are designed around either small hard drives (Creative Labs, Archos) or solid-state memory storage (Pogo Products). Reliability

may be an issue with some of these devices. They nearly always lack a microphone input jack as well as other features that would make them good field recorders. That said, some of these devices have great potential and future developments are worth watching.

MiniDisc (Sharp, Sony)

MiniDisc provides "near CD" quality audio recording, is very portable, has long record times, and is relatively cheap (although the cheapest recorders should be avoided if they lack a microphone jack). MiniDiscs are often used by broadcast journalists and others as a cheap alternative to more expensive field recorders. Disadvantages include a poor computer interface – upload of audio files is only possible by real time re-recording. MiniDisc also needs to be used carefully to ensure directory information is saved or recordings will be lost.

Voice Recorders (Olympus, Sony, Panasonic)

These small solid-state devices are designed to record memos, dictated letters, and the like. Some of the more expensive ones have microphone jacks and interface well with computers through a USB connection or removable flash memory cards. Most of these devices save audio in highly compressed formats, with low sampling frequencies, and limited frequency sensitivity. These factors will limit audio quality. Future models are likely to support higher quality audio.

Professional Solid State Recorders

(Marantz, Denon, Maycom, Mayah, Orban)

These recorders are designed for field recording of interviews by broadcast journalists. They are usually rugged and reliable, have sophisticated recording features, are generally larger than other portable recorders, interface well with computers, and are usually very expensive.

CD-R/RW Recorders

Marantz has recently started to sell a professional portable CD-R/RW recorder (CDR300) designed for recording meetings and interviews. It is expensive but audio quality should be excellent, blank discs are cheap, and audio is easily transferred to computer.

Digital Audio Tape (DAT) Recorders (Sony, Tascam)

DAT is primarily a professional recording medium. It is expensive and is rapidly becoming obsolete.

Personal Note

My own practical experience is presently limited to the use of cassette tape, MiniDisc, and computer recording. I currently make recordings using either a MiniDisc recorder or a computer. In the future, I plan to trade in my MiniDisc for a suitable solid-state recorder. To learn about my approach to using digital audio in my own qualitative research see: <http://www.edc.org/CAEPP/resources/audio.asp>.

References

- Ashmore, Malcolm and Darren Reed 2000 'Innocence and Nostalgia in Conversation Analysis: The Dynamic Relations of Tape and Transcript', *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 1(3). Available at: <http://qualitative-research.net/fqs/fqs-eng.htm>.
- Modaff, J. V. and D. P. Modaff (2000). Technical notes on audio recording. *Research on Language and Social Interaction*. 33 (1), 101-118.
- Muhr, Thomas (2000) "Increasing the Reusability of Qualitative Data with XML", *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 1(3). Available at: <http://qualitative-research.net/fqs/fqs-eng.htm>.
- Patton, Michael 2002 *Qualitative Research and Evaluative Methods*. Third Edition. Thousand Oaks, CA: Sage Publications.

social research UPDATE

(ISSN: 1360-7898)

is published by

Department of Sociology
University of Surrey
Guildford GU2 7XH
United Kingdom.

Tel: 01483 300800

Fax: 01483 689551

Edited by Nigel Gilbert

(e-mail: n.gilbert@soc.surrey.ac.uk)

Autumn 2002 © University of Surrey