

Digital Devices in Sociolinguistic Fieldworks*

[NOTE: This is a largely modified version of Matsuda (2000) and was presented at METHODS (International Conference on Methods in Dialectology) IX Meeting at University of Joensuu on August 5, 2002. In addition to the people I mentioned in Matsuda (2000), I would like to send my cordial thanks to Paul Kerswill, Renée Fürst, T. de Graaf, John Nerbonne, for their comments and discussions at the conference. Renée deserves special thanks here, for providing me with additional references after the conference. It should be stressed, however, that all remaining errors are mine only. This research is partly supported by Grant-in-Aid for Scientific Research #13410139 from Japan Society for the Promotion of Science.]

Kenjiro Matsuda
Kobe Shoin Women's University
kenjiro@icis.shoin.ac.jp

0. Introduction

1. Recyclability of Sociolinguistic Fieldwork Data

With the beginning of Variationist paradigm in 1960s, sociolinguists began tape-recording and analyzing the daily speeches in the speech community. The search for the vernacular (in the sense of Labov 1972) was a clear departure from the traditional dialectological fieldwork that elicits a speaker's response using a questionnaire. The interview may involve one or multiple speakers (group session), but in either case, there is no pre-set time length and the interview may go on as long as the speaker wants to talk. The interviewer may have a course of questions (module) to facilitate the naturalistic interview, so that s/he can obtain a wide range of speech style from the speaker, but the basic principle here is to follow the flow of the talk. Therefore, even if the speaker deviates from the intended response to the question, it is actually a good sign as long as the speaker is spontaneously talking.

A project that involves such a sociolinguistic interview is typically made up with three phases. First comes the interview itself, where the linguist records the whole interview with a tape-recorder. Then the taped interview will be transcribed totally or just scanned for relevant data. The analysis may be about anything from phonetics, phonology to discourse analysis. Finally, the tape will be stored for future use. Note here that under this scheme, unlike the traditional dialectological survey, linguist can always go back to

the original recording to study different phenomenon. For example, an interview tape that was originally intended for phonetic research can be used for a morphological study several years later. In this sense, sociolinguistic interview is *recyclable*. That the data is recyclable and that the data would be used over and over again for different future projects means that the quality of the original recording gains more importance, and the maximum care must be paid to make certain that it is the best quality available and it remains so long after its first production.

2. Some Principles in Fieldwork Data Digitization

Up to 1980s, it was a heavy analog tape-recorder that assured linguists the best quality recordings. But after the digital revolution during the 80s, various digital devices began to hit the general market and attracted linguists' attention. Those new machines were typically smaller than their analog predecessors so that they were more suitable for fieldwork recordings. Furthermore, their price lowered quickly as the revolution proceeded.

The technological progress, however, brought a problem as well. After the debut of compact disk in 1981, it was followed by (excluding different formats of CD) DAT in 1987, Minidisc and DCC (Digital Compact Cassette) in 1991, and the 90s saw a birth of series of DVD formats. Considering the quiet field since 1970s when the audio cassette tape became a default gadget for fieldworkers (with reel-to-reel system still quite active), one could easily imagine that this rush of new formats and devices left many linguists at a loss for what to do with the sudden change in the technological situation. The question that emerges at this point is: How could a sociolinguistic fieldworker benefit from this revolution without going astray?

The cue for solving this problem seems to lie in its very source. Namely, those different recording devices have different characteristics thus each machine/format has its own merits and demerits. Obviously some of them are fit for some specific purposes, while others are not. What is necessary, then, is to find the best match between the purpose, or the phase in the sociolinguistic interview mentioned above, and the currently available digital recording machines.

This implies that the best practice would be the one where different devices are used for different purposes/phases within the same lab or project. While such a configuration may appear inefficient or redundant at first, it is more desirable than the one where a

single device is used for interview, analysis/transcription and storage. This is because the digital technology is equipped with rather rapid metabolism, and there is no guarantee that the currently popular one would survive the next decade. Thus for example, DCC is virtually a museum piece nowadays, and DAT market seems to be quietly shrinking, if not on the verge of extinction. In the worst case, one would then have to face a situation where the media of recording (DAT tape, Minidisc, etc) cannot be played back anymore, if the content was not transferred to a lucky surviving media. In a multiple-device configuration, there is a good chance that one of the devices would survive the selectional process, or at least remain playable just as a CD-DA formatted disk is playable on a DVD player.

Lastly, there must be some consideration of efficiency. It would be meaningless if we came up with a solution that is too complicated and time-consuming, as the time should be spent on analyses in the lab as much as possible. So the possible configuration should also be an efficient one.

Thus, I propose the following three principles as a guideline for digital fieldwork data management:

- I. Survival through diversification
- II. Maximization of merits and minimization of demerits of each available device
- III. The workflow should be constructed in a maximally simple and efficient way.

In the next section, we will see how we could come up with the best configuration under these principles.

3. Comparison of DAT, Minidisc and CD

First of all, let us consider DAT, Minidisc and CD as possible elements in the configuration for the moment. In order to follow the principle II, we need to check the merits and demerits of these three devices. Here I compare the three on six points: data compression (is the audio data compressed before recorded?), rewriting (is the media reusable?), random access (can the media be played back at some random points or should it be sequentially searched?), media durability (does the sound quality deteriorate with time?), recorder size (is it too bulky for fieldwork use?) and media popularity (is it readily available?).

Table 1: Comparison of DAT, Minidisc and CD

	DAT	MINIDISC	CD
DATA COMPRESSION	no	yes (lossy)	No
REWRITING	possible	possible	impossible (CD-R)
RANDOM ACCESS	slow	fast	Fast
MEDIA DURABILITY	limited	very well	Very well
RECORDER SIZE	portable	portable	not portable
AVAILABILITY OF THE MEDIA	no	yes/no	Yes
PLAYABLE ON ...	DAT player	MD player	CD & DVD player PCs

DAT suffers from three points, but its portable size and non-compression recording format are very attractive features of the machine. At the time of this writing, the smallest DAT recorder is Sony TCD-100 (DAT-Walkman) whose size is 80.0mm (W) x 117.3mm (H) x 29.2 (D) mm, so it's readily portable. A CD-recorder is, due to its media size (120mm in diameter compared to 73.0mm x 54.0mm x 10.5mm of a DAT cassette), cannot help but be bulkier than this. The non-compression recording format is also an important point. Minidisc uses ATRAC (*Adaptive Transform Acoustic Coding*) compression system to cut off parts of the original sound, so it is not possible to recover the original input from the recorded one (hence lossy compression). While how seriously ATRAC damages the audio signal for phonetic analyses is still a controversial matter (see §5), DAT safely bypasses this problem by not compressing the signal at all. [NOTE: It is worth noting that MD-to-MD copy builds up compression loss at least for consumer models. This fact alone would be almost enough to make it an undesirable media for the original recording or archiving, both of which would necessarily be copied further in the project.] In view of the data recyclability mentioned above, this is no doubt a desirable characteristics, because the data can be used for acoustic analysis as well as for morphology, syntax, and discourse analysis without any problems.

Except for the possible shortcoming, Minidisc is a fieldworker's dream machine: it has the smallest size of the three [NOTE: the size of the latest model from Sony (MZ-R910) is 80.0mm x 75.5mm x 20.0mm, which is 44% of the DAT recorder (TCD-D100) in volume.], the media is a magneto-optical disk encased in a hard plastic cover that is durable and readily available at local electronic shops at least in Japan and some parts of Europe. Finally, it has a fast random accessibility like CD-R, but it is also rewritable.

The fast random accessibility is an ideal characteristics for transcription or search for relevant tokens.

Then, what about CD? Putting CD-RW (which is rewritable) aside for a moment, CD has several positive characteristics that Minidisc does not have. First, the very fact that it is *non-rewritable* makes it the optimal media for archiving. Secondly, since it does not involve data compression, it is at least desirable for archiving, recording the audio signal as it is. Finally, unlike DVD (see below) its format is highly standardized, and unlike DAT or Minidisc, it is playable not only on CD-players but DVD players and on PCs with a DC drive. Considering the later distribution of the recording to other researchers or data sharing with the project colleagues, such characteristics of the CD is naturally a very attractive one. It is also likely that the media remains at least playable even after the DVD takes the helm of the international audio-media markets. Therefore, even after the birth of a dream machine, CD is still a must tool of the trade for sound-oriented linguists.

4. Configuration of Devices

Taking these characteristics mentioned in the previous section, we would reach the optimal configuration of devices in fieldwork projects as below. Note that each device is connected via digital cables, so the data is copied digitally between the phases.

Table 2: Phase and Device Matching

PHASE	DEVICE
Interview	DAT
Storage / Archiving	CD-R (CD-DA)
Analysis / Transcription	Minidisc PC

There is one more comment that should be made about the table. That is the possibility of feeding the audio data to PC files and conducting all transcriptions and analysis there, especially for acoustic analyses. There are also a number of commercial and free softwares available for that purpose. [NOTE: Praat program by Paul Boersma and David Weenink is one such freeware. Praat is available from: <http://www.fon.hum.uva.nl/praat/>]. From the viewpoint of Principle I, since PC-analysis implies that the data is stored as a computer file, it also contributes to diversification of the storage format, making the data survival more likely.

In term of workflow, the phase-device match suggested above would mean the following schema:

[簡単な図を挿入せよ]

Fig 1: Workflow of Fieldwork Digitization

5. Cons and Pros for the Uses of Minidisc

One might alternatively propose a different model of digitization process in the lab. One such possibility would be the one where all phases are done using the Minidisc. The reason that such model is avoided in favor of the one where three different devices were used is its compression feature ATRAC, which achieves a high compression rate of 5:1 through various psychoacoustic techniques. A question naturally arises then how seriously this operation affects the recorded data, and distorts its acoustic analyses. This issue can be broken into two parts: Compression given to the original data and compression that is made to already compressed data, i.e. cascase compression (Minidisc-to-Minidisc copying). The latter would be the case when the Minidisc is chosen as an archiving format, and another copy of Minidisc is needed (provided that the SCMS code is set to 00 on the archive copy) for transcription or for other purposes.

To begin with latter case, the practice is not recommended at all, as it will only decrease the quality of the succeeding copies (Minidisc Frequently Asked Questions (2002), van Son (2002)). Since an archival copy will be necessarily copied, such a characteristics of Minidisc will bring more trouble to the project than benefits. Thus it follows that the Minidisc is not for an archival use at all.

More controversial than the latter is conducting the interview itself with an Minidisc, a practice that is becoming increasingly popular.[NOTE: For example, Millennium Memory Bank, the biggest European radio and oral history project, adopts the Minidisc as the recording format. See Perks (1998).] There are at least four empirical studies on the subject so far. Kido et al. (1997) compared the recordings through DAT, analog cassette and Minidisc and concluded that at least under good recording conditions, Minidisc recording does not show any apparent acoustic distortions, although they cautiously add that “it is beyond our imagination how the ATRAC will affect the

recordings made under noisy conditions, as often the case with forensic voiceprint analyses” (Kido et al. 1997: 121 -- translated by K.M.). [NOTE: As a member of National Research Institute of Police Science, Kido’s -- as well as Otsuka’s -- research group was interested in how reliable recording evidence made with Minidisc.] The cautious tone becomes more prominent in Otsuka et al. (1998), who conducted spectrographic analyses of synthesized sounds under varying conditions and concluded that it is safe to refrain from using the Minidisc recording unless necessary.

Quite contrary results, however, were reported in Morrison (2001) and van Son (2002). Morrison (2001) is a report of a simple, informal experiment that compared a male voice saying [bit] recorded on to a computer and the one copied to a Minidisc. He found only negligible (0.14%) difference in the F2 reading. Van Son (2002) is probably the latest and the most thorough test on the topic. XXX.

Nevertheless, the current situations surround the Minidisc is controversial at best with cons and pros expressed from various standpoints. Aside from Kido et. al (1997) and Otsuka et al. (1998) mentioned above, Schüller (1999), Matsuda (2000) and Plitchta (2002), are also against using Minidisc as an interviewing tool, or using its recording as the source of acoustic analysis. And even those experimental works suffer from one hitch: they are all based on noise-free laboratory speech. Thus, we still do not know the magnitude of the ATRAC compression effect on recordings made in noisy conditions, but in reality fieldworkers are always confronted with various sources of noises. Until more studies are conducted in this domain, it would be safe to call for a moratorium on the use of Minidisc as a tool for primary recordings.

Concluding, at least until more solid experimental data becomes available, it is safe to limit the use of the Minidisc to a transcription tool for the meantime. Given the rather high evaluation in Table 1, such a position may appear too cautious and too harsh for the dream machine. But at the same time it is important to remember that a given interview could be the last chance to record crucial data for the project. In the case of a speaker of a dying language, for example, it could be the very last speech data available for that language. Then one would naturally try to record the interview without losing *any* part of it, even if that means bringing a slightly bigger machine to the field and dealing with a tape media. Notice that it is not limited to such extreme cases; all sociolinguistic interviews are one-time only encounter, and no speaker will speak in exactly the same manner in the second interview. Rare words, rare phonemes, rare

grammatical constructions --- there is no guarantee that they will occur again in the second interview. [NOTE: One may argue that if the project is of non-phonetic nature, then the use of Minidisc cannot be a problem at all. Actually it *is* a problem, because, as mentioned above, sociolinguistic data is recyclable. Schüller (1999) also argues from the viewpoint of oral history: “Although primarily aimed to capture the mere content of narratives, it should be kept in mind that oral history recordings are important socio-linguistic sources, which will give future generations of researchers and excellent insight into how we speak, pronounce, and phrase in our daily use of the language. Hence, even if elaborate evaluation such as sound analysis is not the purpose of such projects today, and may presently seem beyond the scope of the further use of the recorded material, it can be taken for granted that such recordings will in future also be of great value to generations of linguists, psychologists, and others. Such future users will be most grateful to have an audio signal which will best serve their aims”.] In light of such nature of sociolinguistic interview, our cautious attitude may be well justified.

6. SCMS

Any discussion of digital audio data handling is not complete without mentioning SCMS (Serial Copy Management System). Originally developed to prevent infinite digital copying, the SCMS sometimes works as a trap for field linguists. The basic principle of SCMS is that it prohibits one from making a digital grandchild, while allowing infinite analog copying. This situation is depicted in Fig 2 below:

Fig. 2: SCMS
[METHODS の図を引け]

Probably the most crucial point of SCMS is that it applies to any digital data, *irrespective of who recorded it*, as long as it is used on non-professional machines. [NOTE: Actually if the SCMS code was set 00 at the time of recording, it would permit further digital copying. But such coding is only possible on professional recorders.] What it means is that in the worst case, one may not be able to digital-copy the fieldwork recording s/he made before; the original DAT tape was copied to the archiving CD-R, and then the CD-R cannot be copied further. So if it were the only survivor, no digital copying would be possible.

There are several things one could do to cope with the situation. First, the SCMS only applies to commercial models, so it is a good idea to have at least one professional

machine in the lab to override the restriction. Secondly, SCSM does not restrict the number of copies made from the same source, so it is advisable to make a Minidisc copy for transcription and the archival CD-R copy from the same DAT tape, instead of attempting to copy in a serial manner as DAT-> CD-R -> Minidisc. Lastly, since SCMS is a restriction only on digital copying, there is always a possibility of analog copying.

7. Final Remarks

Lastly but not least at all, some words about DVD is in order here. DVD is of course one of the latest item in the digital revolution and with its gigantic size of 4.7GB, surpassed the VHS in the visual recording sector. Then, why did not we mention it in this report? The problem of DVD is that its format has never been standardized, resulting in the nightmare of numerous different formats supported by different companies -- DVD-RAM, DVD-R, DVD-RW and DVD+RW. This situation is in stark contrast with compact disk, whose format was successfully standardized and enabled researchers to exchange data across different machines. Thus, an interview recording stored in a CD-DA format is now playable on any CD-players and almost all PCs with a CD drive anywhere in the world. Such an ideal state for data exchange is a far cry for the DVD, and there does not seem to any prominent move for standardization. Probably the best thing to do is to use it as an auxiliary archive media for backup.

Concluding, the digital revolution during the 1980s was definitely bliss for fieldwork-oriented linguists. But it is quite easy to get lost in the flood of new gadgets and misleading information. As a response to this situation, I proposed three guiding principles in digitization of fieldwork data and the DAT/CD-R/Minidisc configuration. The Minidisc has a good potential to be a true dream machine for field linguists, but at least as an interview recorder, it is advisable to call for a moratorium until more empirical evidence is available regarding the degree with which the ATRAC affects acoustic properties of the audio signal. Given the increasing popularity of the machine in Japan and Europe, such reports are well overdue and will certainly be beneficial to the community of field linguists.

References

Kido, Hiroshi, Weizhong Zhu, Takashi Nagauchi, Toshiaki Kamada, Masumi Tanimoto and Hideki Kasuya. 1997. Seimon kantei-ni okeru Minidisc rokuon-no tekihi-ni tsuite ("On the usability of Minidisc-recordings for voiceprint analyses"). *Nihon Kanshiki Kagaku Gijutsu Gakkai Dai3-kai Gakujutsu Kôen Yôshishû*. 121.

- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Matsuda, Kenjiro. 2000. Fiirodowaaku Shiryô-no Dezitaru-ka ("Digitizing fieldwork data"). *Theoretical and Applied Linguistics at Kobe Shoin*. No. 3 27-45.
- Minidisc Frequently Asked Questions*. 2002. http://www.minidisc.org/minidisc_faq.html
- Morrison, Geoff. 2001. Posting on Yahoo! Group "Praat Users Group" #204 (Oct. 17, 2001). <http://uk.groups.yahoo.com/group/praat-users/message/335>
- Otsuka, Norio, Ichiro Hayama, Nagauchi Takashi, Tanimoto Masumi. 1998. **MiniDisc rokuon-ni okeru supekutoru-no rekka-ni tsuite** ("On the spectrographic deterioration in MiniDisc recordings"). *Nihon Kanshiki Kagaku Gijutsu Gakkai Dai4-kai Gakujutsu Kôen Yôshishû*. 95.
- Perks, Rob. 1998. "Millenium Memory Bank." *IASA Information Bulletin*. No. 27. 12-13.
- Plitchta, Bartek. 2002. "Best practices in the acquisition, processing, and analysis of acoustic speech signals." *Handout for METHODS XI Workshop "Recording and Analyzing Linguistic Data."* Joensuu: University of Joensuu. August 5, 2002.
- Schüller, Dietrich. 1999. Minidisc for field recording? Applying archiving principles to data gathering. *IASA Journal* No 14. 35-40.
- Van Son, R.J.J.H. 2002. Can standard analysis tools be used on decompressed speech? Paper presented at CoCOSDA Workshop. Denver. CO.