

# Using live text transcription for qualitative research interviews: methods notes after a first experiment

<http://blog.melchua.com/2013/01/23/using-live-text-transcription-for-qualitative-research-interviews-methods-notes-after-a-first-experiment/>

May 8, 2013

January 23, 2013 – 5:18 pm

This is a research journal note on using live text transcription (via CART or other means) for qualitative research interviews. I'm typing in the Glenn building on OSU's campus; Kyler is next to me reading my cultural theories reader (by Glesne) with a concerned look on his face (it's a hard book) and we've just both come from a mind-spinning scholarly autobiographical talk by Patti Lather. So much for scene-setting. I'm also typing these initial paragraphs to warm up, mentally and physically — it's cold and my fingers are defrosting.

Here's what I've got. The "normal" interview protocol is something like this:

## Current interview protocol

1. I (the researcher) talk with the subject and tape-record our conversation.
2. I go home and get the tape transcribed (delay!)
3. I send the subject the transcript and say "could you check if that's right?" (delay!)
4. Busy subject does not have time to read long transcript of a conversation he/she no longer remembers (delay!)
5. I try repeatedly to contact subject until (1) I give up, (2) subject says "stop bugging me," or (3) subject caves and checks the transcript (delay!)

Most researchers who use interviews do all these steps. Since I'm trying to practice [radically transparent research](#), I add another step after all these: if the subject grants permission to release the edited/revised data (maybe anonymizing names or taking out some parts) under an open license, then the data enters an open dataset, and analysis on it can also be done publicly, and other fun things. (This is instead of the "normal" practice of having the data be in a secret place that nobody can see, so people have to trust the researchers to have interpreted it "correctly" and they can't reuse that data for other things — in open source software terms, it's like releasing a binary blob.)

You'll notice there are many delays in the above process. You need to wait for transcripts, then wait for the subject to see their transcript, then wait for... and every waiting moment increases the chance of participant dropout.

Here's what I want the process to look like instead.

## Interview protocol with realtime transcription

1. Talk with the subject and have our convo transcribed in realtime and displayed on a screen we both can see. (Probably remote CART with a tablet to the side where both subject and I can see it.)
2. Immediately after — or even during — the interview, I tell the subject "ok, go and edit/cut

whatever you don't want public, and tell me when you've got a version you'd like to apply an open license to and publish."

3. They do that. We push a button. Bam, open-licensed data is available to everyone — immediately after it's generated.
4. There is no next step. We are already done. I don't need to juggle follow-ups. The subject is free and clear.

That's the theory. How does this look in practice? To find out, I tried two things: piloting that interview technique (to see what parts of my hypothesizing fall apart in reality) and asking my friend, NYC CART provider [Mirabai Knight](#), for her thoughts. Here's what happened, summarized by theme.

**Worry: people get awkward and self-conscious when they see their speech being transcribed.**

Mirabai mentioned that some people might "feel self-conscious and clam up and get distracted reading what they just said, so some of your subjects might request to have the screen pointed away from them, so it doesn't throw them off." This actually didn't matter in the pilot. It did take a few minutes to explain the setup (since it's different from what most people have seen before), but once we started talking, both my subject and I just ignored the screen. This may have been because the transcript lagged a few seconds behind our actual conversation, so we had to pay attention to each other to keep up anyway — but in any case, it didn't interfere with our conversational process. If it had, we could have moved the screen anywhere the subject wanted it.

Also, each subject will get interviewed multiple times (at least for my dissertation, and for many qualitative research projects) so that "learning curve" is really a first-time setup thing, and on later interviews it'll just be "ah right, that's part of the way we do these interviews."

**Worry: if we read the screen, it'll interfere with eye contact and therefore rapport. If we don't, we'll miss incorrect transcriptions when they come up.**

This was another tradeoff brought up by Mirabai, and one I ended up being overconfident about. "I read text extremely fast and lipread extremely well and tend to be very, very good at patching different input streams together without losing a connection with the person I'm talking with," I told her. "I think this will be ok."

And it was ok — but because I ended up ignoring the screen and lipreading my subject, as mentioned earlier. Looking at the screen does noticeably break eye contact, and this felt like a rapport diminisher — in contrast to note-taking, which can be a rapport-*builder* even if you need to interrupt your subject to do it. "Hang on, that sounded really important; I want to make sure we write that down." Maybe that's because of the active nature of the writing/typing of notes; the subject can see you're *doing* something in response to them, marking their words as Serious Data — in contrast, the flickering of eyes to screen looks a lot more like "I am not paying attention."

That's something I want to play with in the next round; is there a way I can make my (our?) engagements with the transcript more visibly *work*, more visibly "this is because I am paying *more* attention to the important things you're saying," more active, more engaging?

We're also still left with the other part of the problem Mirabai pointed out, that of missing mistakes: "...misheard phrase or a misspelled proper name or if the subject was talking too quickly so the CART provider had to condense, or maybe even made a misstroke and didn't catch it..."

For the next pilot experimental round, I want to try solving both those problems with the same

technique: instead of trying to edit the manuscript into perfection in realtime, let's just play with the transcript together immediately after the interview and check it then *for content edits* — that the realtime transcript will become valuable during what I'm going to call the *immediate participant check*.

### **Working out the immediate participant check**

Mirabai, again: "If you're mainly relying on the subject to edit and correct things after the fact, that would probably work quite well in most circumstances, especially for fairly short interviews. If it was a long interview and there's a lot of text for them to wade through, they might be too worn out to proofread carefully."

This is an excellent incentive to do short interviews, which I should be aiming for anyway. I'll need to figure out what the steady-state storytelling time is for most people — that is, can I get people to practice telling (different) stories from their lives until they can fit each story they tell under a certain time-goal (say, half an hour?) If so, how long does it take someone to "train" into the half-hour (or however long the time-goal is) format, so that even new stories come out in that conciseness and length?

Mirabai's comment also made me realize I needed to decide what to do and what to *not* do during immediate participant checks. Going for 100% realtime transcription accuracy was already abandoned as an unreasonable goal. Going for 100% transcript review during immediate participant checks... also probably unreasonable. I'll need to distinguish between *transcription errors* (misspellings, misstrokes, phrases transcribed a little off, or condensed) which we can note but do not have to fix during the immediate participant check (I can go back with a videorecording of the interview that I can lipread and do those fine-detail fixes later) and *content edits*, which are what I really want the participant to do: "Oh, make sure you anonymize this part — and take that section out."

What I want to leave with is *not* a finished transcript ready for release, but a to-do list from the participant on what to *do* to the transcript so it will be finished and ready for release — an agreement that if I do X, Y, and Z, then their data is free to go (without needing to check in with them further). For instance, "make sure that name is spelled right through the whole transcript, and delete the tape from here to here, and let's change this person's name to Joe, and just say that they're from a small African village" might be one set of instructions.

I'll need to make sure my interviewees don't feel rushed. Mirabai pointed out that they might feel awkward about their reading speed in front of me (feeling like they're wasting my time unless they review it as quickly as possible), or that they might want to sleep on a particularly sensitive bit of information before deciding whether they're comfortable with releasing it. I want to see if I can make transcripts more navigable to help alleviate some of these issues, and have a software engineering proposal (which I'll write up later) as to how this could be the case — but it won't remove these issues entirely, and right now all I can say is "I need to pilot more and be conscious and watchful of this; it is not yet figured out."

That's what I have for this round. My next round of notes on this are likely to be about liability concerns, technical suggestions for how I could make transcripts more navigable during immediate participant checks, and how this method of transcription also brings with it some default philosophical stances about data that I should probably point out explicitly when I'm writing this up for my prelim.