

Title: Graphic Transcripts

Abstract: In this entry I introduce graphic transcripts as a hybrid form between comic strips and transcripts of practices, which are used in discourse studies. The basic components of panels, gutters and speech bubbles etc. are explained. Graphic transcripts are then compared to textual transcript formats, especially the Jeffersonian format. I provide examples of studies where graphic transcripts have been used and the reasons for their use. In closing I reflect on the tension between establishing conventions for graphic transcripts and the need to tailor transcribing for the particular thing the researcher is investigating.

Keywords: transcription; graphic transcription; comic strips; interaction; visibility

Author: Eric Laurier

Address: School of GeoSciences, Drummond Street, The University of Edinburgh, Edinburgh, EH8 9XP

eric.laurier@ed.ac.uk

<https://orcid.org/0000-0001-8043-0710>

6500 words

Graphic Transcripts: convention, production and showing

Graphic transcripts are a way of transcribing recordings of discursive practices which borrow from comic strips familiar formats for representing speech and other actions. In borrowing the formats they have been adapted to the particular purposes and criteria of transcription. Thereby the graphic transcript has become another of the serious cousins, like the graphic novel, which are part of the sprawling family of the comic strip. Their purposes, in common with representations of empirical material elsewhere in discourse research, are to provide readers with shared access to those discursive materials, as forms of data, evidence, fieldwork stories, vignettes and so on.

As part of the emergence of conversation analysis in the 1960s, Gail Jefferson revolutionised transcription by altering the conventions for representing audio recordings. To achieve this, she experimented with the technology of the time which was the mechanical typewriter, the pen and paper sheet. Her background was not in phonetics, instead it was from sociology, a discipline which transcribed talk into the kind of dialogue familiar from newspaper interviews and theatre scripts. What Jefferson had in mind was rendering the details of talk which were being examined by herself and her colleagues Harvey Sacks and Emanuel Schegloff. Details which although hearable, were ignored, corrected or altered in representing speech in sociology. One example which she provided as a case for transcribing what was heard, was the 'nyem', it was something the conversationalists she was listening to were saying and hearing. It could have been corrected to a 'no', 'yes' or an 'ehm', yet the very point was that the speakers were using it to do different things that saying either 'yes', 'no' or 'ehm' (Jefferson 1978). There were many other hearable and accountable features for which she created typewriter-based markings such as overlapping speech, accelerated and slowed speech which have become the basis of a standardised transcription system used by researchers studying talk-in-interaction (Hepburn & Bolden 2017).

While new audio recording technologies provided the means and occasion for researchers to hear and account for details of how we interact that were otherwise heard but un-noticed, they led to other problems. The most significant is captured in Bogen's complaint and his complaint is two-fold (Bogen 1999). First, the art of narration has been displaced by a mechanical reproduction of recorded details, where the transcriber should be trying to retell what happened in the event they are instead transferring *all* that is hearable on the recording. As a result, the transcription's detail is both flattening and overwhelming for the reader. Second and relatedly, the reader becomes reliant on the author to retrieve, from the detail, what is significant. The accompanying text around the transcript is directed toward making sense of the transcript at the expense of description, analysis and argument. Bogen's complaint about Jeffersonian transcripts is one to keep in mind, now and later, as a broad principle for readability in producing transcripts be they textual or graphic.

Borrowings and adaptations

As I noted earlier, the graphic transcript is a variant of the comic strip and borrows its basic conventions. It is constructed out of a sequence of panels with images which have gutters around them; readers read across the panels re-animating them into a flow of action. There are other familiar components, though they are not always all used in transcribing: speech bubbles, thought bubbles, caption boxes, motion lines, emanata (symbols which emanate from a character indicating what that character is feeling), background sounds and more. Although we usually read the panels left to right, just as we would read the text of a sentence, comics support different reading practices and are read in various ways. For example, they may have a large

central splash page, be more diagrammatic or borrow from other styles which support readings which rove around the page (for 99 different possibilities, see Madden 2005). Their layout provides greater flexibility than the conventions of line-by-line textual transcripts, a quality which brings with it both interesting possibilities and abiding problems.

As researchers of discourse, when our task is to create transcripts out of recordings, what are the uses which we can make of comic strip conventions? Most obviously, we can attach *speech bubbles* to whichever party says those words. In each panel, we can include a screen grab from the video recording, or a drawing of a screen grab, as the *panel image* of what is happening. *Captions* offer us the options of: providing additional information; characterising what is happening, resolving ambiguities in the screen-grab and more. *Thought bubbles*? In my experience, transcribers have included them in their transcripts to provide their interpretation of what a person is thinking, yet not saying aloud. While this is a mistake by the criteria for transcribing, it is a useful moment to ask yourself to consider: what is it in the observable and reportable actions of the party which led you to add a thought bubble. In other words, how could you reasonably surmise what they were thinking? It might have been a moment where someone was expected to say something aloud and instead there was a noticeable silence. It might be that they make a facial expression in response to something being done to them, one which is easily intelligible in its own terms, though not obvious when shown as a screen grab. As a practical solution to the latter, there is the option to use a caption box to help the reader see what you witnessed in the recording.

Having briefly considered what familiar formats from comic strips we can borrow, what elements of transcribing do comic strips *not* represent? These absences echo those that Jefferson alerted us to in existing textual transcripts of speech. One feature is the common phenomena which Jefferson successfully documented: *overlapping speech*. For the the graphic transcript overlapping speech can be depicted by overlapping speech bubbles (see figure 1, panel 2). They lack the precision of Jeffersonian markings, where, of course, just when, in talk, overlapping talk begins is informative for us of just what the overlapping speech was doing: interrupting, agreeing, acknowledging and so on. Marking overlapping talk is not insurmountable as a design concern and how carefully it is represented will depend on what it is that you wish to ultimately show in your transcript.

Missing speakers. Because we are using screen grabs from recordings, our speakers are sometimes not in view of the camera (figure 1, panel 3) and so are missing from the picture which is not a problem for comic strip artists, instead for them it is a storytelling device. Arising from our twin concern with the *sequencing* and representing of what is happening, the depiction of changing actions, perspectives, appearances of new objects and persons can require a second panel where the speech is continuing across these panels (fig 1, panel 4). There is a satisfying simplicity to showing talk where a *continuer* or *other particles* are said (fig 1, panel 5).

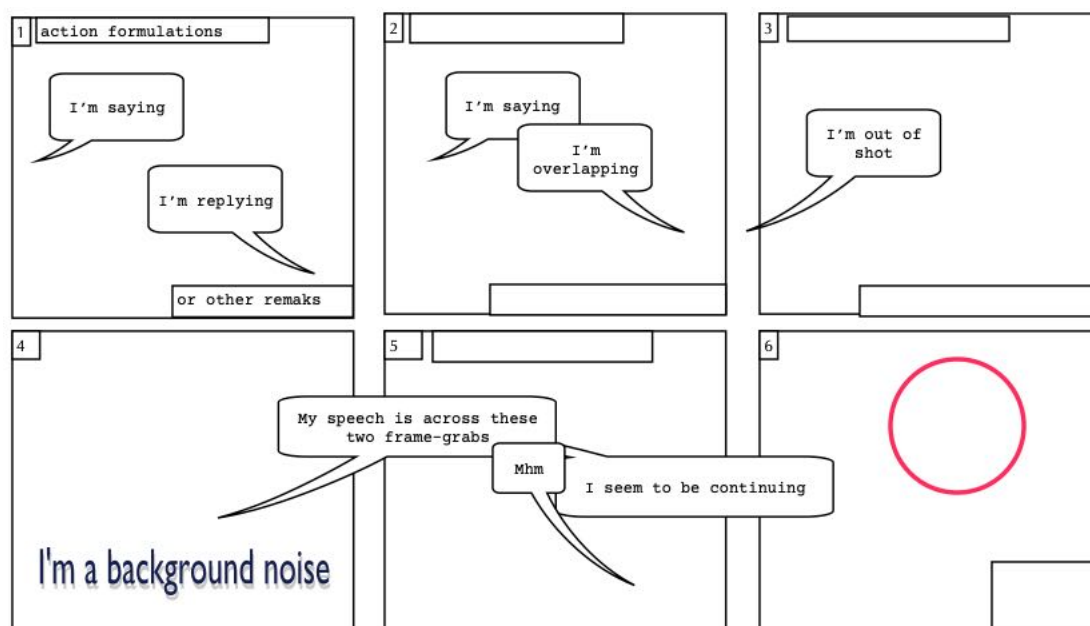


Figure 1 – Graphic Transcript Template 1 (source: author)

Because transcripts sit within talk at data sessions, or are part of texts, both of which settings provide accounts and analyses of them, they have *line numbers* and *annotations*. Line numbers in graphic transcripts become panel numbers. While transcript *annotation* is not unknown in comic strips, it is unusual. There are several iconic shapes which are routinely used for highlighting features in transcripts. In my template, it is a red circle (fig 1, panel 6). Arrows are commonplace too.

Beyond these basics illustrated in figure 1, the comic form offers rich and complex possibilities for transcribing which I have explored in earlier publications (Laurier 2014; Laurier 2019; Laurier & Boelt Back 2023). While the panels in figures 1 and 2 have set widths, we can change the width of panels to achieve various effects. Width is used for allowing for a wider or narrower scenes to be drawn; of interest in depicting action, it can also be used to provide a sense of *duration* and *rhythm* (McCloud 1993). Width, for the transcriber of actions, provides a useful resource for indicating the slowness or fastness, and, brevity or extensions of what is represented panel by panel (see figure 2, panels 1 and 2).

As I noted earlier, one of Jefferson's breakthroughs in representing speech was using square brackets (hand drawn in early transcripts) to mark overlapping speech. The reader has the textual depiction of hearing two things being said at the same time and for the graphic transcript these may also be visually recognisable and visually organised courses of action. For these *overlapping courses of action*, removing the gutter between two panels is a way of indicating their simultaneity (figure 2, panel 4a and 4b).

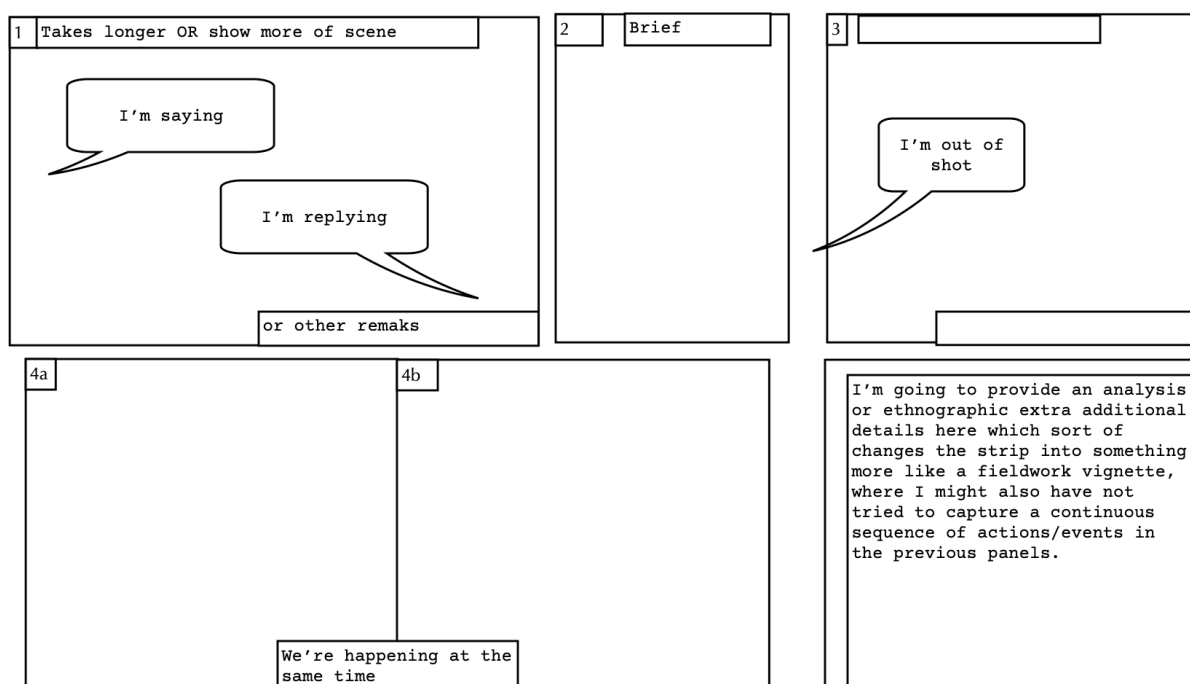


Figure 2 – Graphic Transcript Template 2 (source: author)

Studies where graphics transcripts are particularly appropriate

There are a number of examples we can turn to which demonstrate discursive practices where using graphic transcripts is particularly salient and fruitful. They range across a wide set of discursive settings where visual materials are significant to those where there is an absence of hearable talk.

In my own research on mobility and interaction, graphic transcripts provide visual access to the setting in which people are moving. For example, representing what features of the road drivers and passengers look around for, and see, and then respond to by slowing down, speeding up and honking their car horn. In terms of visual media, it could be, for pedestrians, or car travellers, navigating using a map app. As investigators, we need to have access to what the navigators are seeing on their map when they turn themselves around or move forward (Laurier et al. 2016). Without access to the visual objects they are examining and what is happening in their visual environment, we would struggle to understand the sense of their talk. In a different vein, there are workplace sites where visual materials are central to doing the job such as in architectural practice where action takes place around maps, diagrams, building plans and materials (Buscher 2006; Ivarsson 2010).

In understanding classroom discourse the graphic transcript can be utilised to show just what teachers are showing and pupils are inspecting. In fact, there have been a series of studies by Wallner (e.g. 2017) where the classroom work is itself comic strip making. Graphic transcripts lend themselves to studies of bodily displayed understanding or lack of understanding by students. In Tyler's (2021) study of 'whole body sensemaking' she accomplishes this in transcripts that are hand drawn from screengrabs. Rather than draw them for herself she collaborated with an illustrator which is a way of working that I have also found inspiring. It requires familiarising the illustrator with your research approach and with the criteria for transcripts.

Graphic transcripts have been used in studies seeking to understand everyday interactions with robots (figure 3), where robots appear to make eye contact with humans and consequently gaze is transcribed simply by connecting arrows between the robot and others participants in the interaction (Pelikan & Hofstetter 2023a; Pelikan and Hofstetter 2023b). Where embodied aspects of eating and caring are the primary focus, graphic transcripts have been used without recourse to multimodal transcripts, thereby foregrounding the practices of eating with very young children whose speaking remains limited (Wiggins et al. 2024).

Extract 3a: Cozmo at the dinner table. RobotCorpus-A2. Translated from German. Transcription symbols: $\square$ overlapping utterances, .h inbreath, ::: prolonged sound. Numerals indicate order of events. See Jefferson/Mondada transcript in A.3.

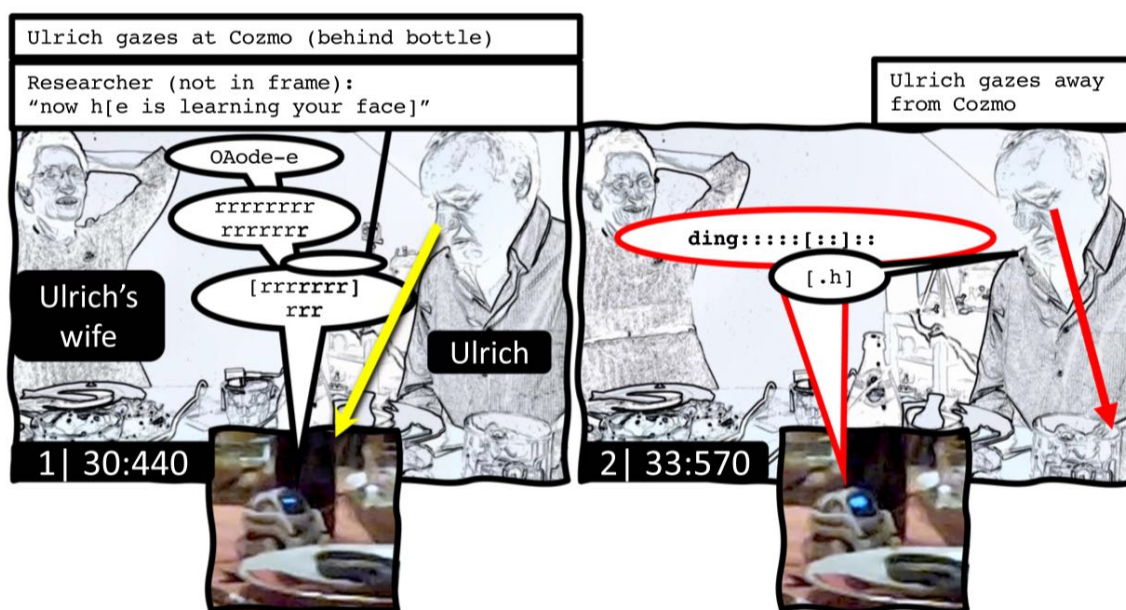


Figure 3: human robot interaction (source: Pelikan & Hofstetter 2023b)

Using graphic transcripts to describe training in mountain rescue training Smith ties their use to analytic issues which go beyond those of visual access noting, that graphic transcripts avoid losing ‘the gestalt contexture of the scene’ (Smith 2024:62). It is not only then that more of the situation is offered to the reader for their consideration, it is also that by transcripts presenting an image of a person rather than labelling a person avoids the ‘single-category labelling’ (ibid.) of those actors in the transcript.

A final example seems one where the graphic transcript would solve an abiding problem – signing conversations. In studying Norwegian Sign Language, Skedsmo’s (2021) chose graphic transcripts, because there is no audible talk and the conversations are visually produced and recognised and organised by their members. As Skedsmo notes, signed language in its details is typically glossed through the process of transcribing into written equivalents which retain something of its embodied qualities or fully translated as if it were audible spoken talk. While it might seem then that a graphic transcript would then support showing the signing, Skedsmo shows that the level of granularity required would make for very long transcripts, though, as an aside, he notes that there is a deaf Marvel Avenger character who signs. Skedsmo’s solution to

the problem of representing signing was to adapt speech bubbles to provide English translations for non-speakers of Norwegian sign language.

Across these studies we find a series of practices where graphic transcripts have lent themselves to transcribing where people and things are, phenomena which they are looking at, whether they are looking at one another (or appear to be, in the case of robots) and the appearances of various visual phenomena. In sum, studies where the situation is one which involves many more courses of action than audible talk.

Comparing Graphic Transcripts with Multi-Modal Transcripts

In deciding whether to use graphic transcripts or not, what is beneficial is to have a sense of both what benefits they bring and what limits they have, particularly in relation to continuing with textually organised transcripts. The main format for comparison with graphic transcripts is the multimodal format formalised in large part by Lorenza Mondada (2018). She has led the way in augmenting the Jeffersonian format for rendering embodied actions and her work has sought to bring a set of standards inherited from a wider and wilder array of video transcription practices from the previous decades of transcribing video. Many of these earlier approaches inserted images as mere illustrations, subordinate to the textual transcript (Ayaß 2015). Alongside Mondada, Paul Luff, Christian Heath and their colleagues pioneered transcription practices which gave the visual organisation of discourse primary importance in their transcription (Luff and Heath 2015) as did the group around Candy and Chuck Goodwin. A recent history of the this latter group has a wonderful proto-graphic transcript of one party lighting the cigarette for another without words (Harness Goodwin & Cekaite 2025).

Our first issue in weighing up graphic transcripts is their representation of the timing of speech and other actions. The relative timing of actions is a central sense-making resource in talk-in-interaction which the Mondada formats seeks to preserve. The mono-space font transcripts which are the dominant form in transcribing talk and indeed in texts for theatre, film, and television, provides just such a sense of timing. With timing at the heart of a transcript, it comes close to a music score. Indeed, early analysts of video recordings experimented with timelines stretching across graph papers which were very close to musical scores (Luff and Heath 2015). By comparison, the graphic transcript has no one timeline to have actions pinned on to. Instead, it has the multiple temporalities provided by: speech in the speech bubble; the image in each panel; and, potentially those of a caption, if it contains further indications of the temporality of what is happening. In figure 4's first panel there are three timings: the timing of the caption 'C1 flips purse to look for change', the sense of duration from the image itself and the timing of the talk 'I might have Jenny'. In the second row of the transcript there is a fourth sense of timing created by the panel widths where both the rustling of the paper money and the customer's nod have a sense of the brevity of the action happening at that point in the sequence.



Figure 4: Graphic transcript readable with multiple temporalities (source: Laurier 2019)

There are further alterations that that can be made to graphic transcripts, in order to preserve timing by subordinating panel sequences to a timeline. Figure 5 was a format developed in collaboration with the illustrator Ryan Hamill which hybridised Jeffersonian with comic strips features.



Figure 5 – Hybrid graphic transcript built around line-by-line speech, readable with single timeline (source: Laurier 2019)

What might appear to be an insignificant matter in transcribing is our second issue: the *categorisation* of each speaker for labelling their turns in the transcript. These categorisations are proto-analyses of what is happening in the recording which push us toward staying with one identity for these persons throughout the transcript (Smith 2024). In figures 4 and 5 these are ‘C1’ and ‘C2’ which then stand for customer 1 and customer 2. The reason they have these labels

in figure 4 is because in this hybrid between Jeffersonian format and comic strips, speaker labelling is required and figure 5 they would have been best removed. A label is itself a proto-analysis of who the person is speaking as. By labelling the members of the setting as customers, we risk missing both their shifting identities which are made relevant by self and other categorisations, changing activities and, moreover, one of their potential identities for each other in this talk, as friends, not simply customers, sorting out a paying problem. Friends potentially sorting out such a problem in distinct ways and with distinct accountabilities compared to two unacquainted customers. In a more conventional comic strip format, speakers are not categorised by the analyst. As we saw in figure 1 it is the tail of the speech bubble indicates which party is speaking. Through what identity they are acting as remains to be determined in considering just what they are doing.

A third issue in relation to graphic transcripts, which grows out of their possibility of supporting multiple temporalities, is *granularity* (Lindwall & Lymer 2024). By granularity I mean the level of detail in which we transcribe what is happening. Note that granularity is not to be confused with the social sciences pair of macro and micro which are their own migraine-inducing contrast for any scholar of social action. Granularity is the level of detail produced in the transcript. Detailing itself is an infinite task. For example, in figure 4 we could have broken down the opening of the purse and extracting of money into a 12 panel sequence or shortened it just 1 panel. While being detailed is a key concern for transcribing, how detailed to be can be led by how detailed the participants are in their practices. As Schegloff puts it: ‘the level of granularity at which noticing is done matters not only for the social actors being studied, but for us as investigators as well; so too at what level the observed or noticed world is described’ (2000: 719)

Perhaps a final issue and one which brings us back to Jefferson’s early concerns around caricature and it’s notable that many of the semi-phonetic transcriptions she experimented with have dropped away because they ‘invoke an images of deviance that project onto the speakers represented in the transcript’ (Murphy 2021: 4). The legacy of the comic strip was one of fun, subcultures and a reader who was a child, which the very term ‘graphic novel’ was used to mark a response to that legacy.

Graphic Transcripts as Work-spaces

Broadly, I would not recommend using graphic transcripts as workspaces, the simplicity of a printed and usually simplified Jeffersonian transcript for annotating and note-taking is hard to beat. A timing-based transcript provides a positioning and running order of what happens that allows you to recover your noticings in relations parts of your original recording. These transcripts can be augmented through using software systems that link transcripts exactly to video recordings. Many people find DOTE, Transana and similar apps essential to how they work. You will have to weigh up the labour costs and your preferences for ways of working.

That said, new technological possibilities are shifting what we can do in terms of the drawing of, and thereby the dawning of, our ideas when we are watching recordings (Albert et al. 2019). They hark back to the earliest days of transcribing video recording, where transcribers experimented with tracing images using paper layered on top of the video screen. Moreover, there are ways of moving transcribing to a stage further forward in the analysis. Instead of producing transcripts in advance of replaying recordings, where the analysis is positioned as second stage which is being annotated on top of the transcript, the researcher(s) are provided with an empty set of 3 or 6 panels. Why 3 to 6 panels? It could be 2, just to show paired or mis-paired actions. Once there are 3 then there are many more practices to document in that way:

initiation + action + completion; action + response + repair; trouble + assistance offer + refusal etc.

In watching and re-watching the recording, when the analyst(s) notice a phenomenon of interest in the recording they attempt to document it in 3 or more screen grabs. These 3 to 6 panel strips can then be talked around if it is a group data session or further annotated if the researcher is working alone. It is a different way of getting started that foregrounds the embodied and visual aspects of a setting (for a longer description of this process, see Laurier & Boelt Back 2023).

Graphic Transcripts as Showing the Thing

As Ayaß says of multimodal transcripts, the ‘high level of complexity runs the risk of making the transcript difficult to read. To understand it, readers must first decipher it, which is at times rather tedious and may conflict with the maxim of readability’. (Ayaß 2015: 519). Keith Murphy, who produces well composed and appealing transcripts, goes further and complains that most published transcripts are bad, ‘technically bad’ (Murphy 2021: 2). A complaint that would, I suspect, apply to many of my early attempts at graphic transcripts which were uninviting and often just plain hard work to understand what was happening in them. The bad design comes from the process, which as Murphy points out, remains with detailing the recording rather than moving forwards to showing the phenomenon.

‘most reflexive attention to how events, recordings, and transcripts relate has looked backward, focusing on how the transcript corresponds with a previously recorded event ... rather than forward toward how the transcript will circulate and be used as an analytic instrument.’ (Murphy 2021: 2)

Prioritising aesthetic and design criteria, Murphy proposes four principles to follow in producing transcripts. The primary move, and most obvious, which does not make it the easiest, for the very reason that we have been so heavily trained in the handling and layout of ASCII text, is a shift to ‘graphic process’ (Murphy 2021: 3). In shifting our attention and effort to an aesthetic of visual showing, is to ‘work compositionally’. Instead of visual elements being fixed, they are to be treated as malleable, modifiable and calibratable. Text (which is also visual, of course) can be made to convey more by, for example, changing its fonts to indicate characteristics such as hearable emotion. Images can be line-drawn to pick out features of interest, most often the outlines of people and the objects that are using. Murphy makes clear, through examples, that there several possibilities for organising items on the page which display them in orders of relevance or sequence. Colours, too often overlooked, allow for a matching of text with speakers or objects. Bringing us back to our focus in this chapter, Murphy stresses the importance of where and how images are laid out in relation to text, where the graphic transcript is taken to be one broad domain of application of the principles he is suggesting.

A first application of his principles to graphic transcripts is in the selection of fonts. All too many researchers, when trying out graphic transcripts for the first time, use the default (and fun) font which is usually supplied with a comic making app. A fun font, which as noted above, is usually targeting children aged 12 and under, is usually a hand-writing font. It is style to consider for a graphic novel aesthetic but there are good reasons for leaving the fun font behind and hybridising with existing transcription fonts. Monospace fonts bring a familiar set of augmentations for representing both timing and hearable qualities of speech. Murphy reminds us that the Courier font (which is the one requires for contributors to this book) has more to do with the available technology when Jefferson and others were developing transcription formats

and no longer needs not be the default choice for published transcripts. While I also continue to default to Courier, I have swapped to Courier Prime which is a better weighted version of Courier New (the latter is pre-installed on most digital devices).

The selection of the screen grabs which fill each panel and stand as icons, poses analytic and representational issues (Ayaß 2015; Stukenbrock 2009). Each panel image stands for both a moment in time, a snapshot, while also representing an action which will inescapably be seen as having a duration. As readers of comic strips we read the duration into the individual and sequenced representations of actions. For example, a single panel of someone drinking a cup of coffee, in-between two panels of them cleaning the house, is one that we read as having a duration established both by our sense of coffee breaks and from the activities it is taking a break from. If drinking and eating as activities are, themselves, our interests, we might be capturing the moment that the action of lifting the cup begins, the moment it meets the drinker's mouth and then whether it hovers or not, or if it is returned promptly to the table. In figure 5, each screengrab is selected to show that there were hands reaching out to comfort a passenger, not selected for when reached out was initiated and when they were retracted. There was nevertheless a temporality to the reaching out because these comforting hands reached out after racist language had been used by the woman pictured and remained in place across several panels (Laurier et al. 2024).

Anonymisation of images in sensitive data often produce problems for retaining identifying details of what is happening. It is difficult to find anonymisation filters which do not make details of faces, texts and objects harder to see and will require trying out different ones. Pencil sketch, paint or 'comic' filters, usually part of comic making apps, are good places to start and the possibilities for tailored alteration of images is constantly increasing. There are though also solutions to the ambiguities and opaqueness created by filters, for example, in figure 5 it hands that were reaching across to console a passenger which are re-established through both annotation and captioning. Meantime, figure 6 also illustrates how speech bubbles and captions can be placed for additional anonymisation purposes over the face of a young child. Other solutions to anonymising graphic transcripts than using filters, are to re-draw, correct by hand or trace over the original screengrabs.

Figure 6: Anonymisation filter, annotation and captions (source: Laurier et. al 2024)

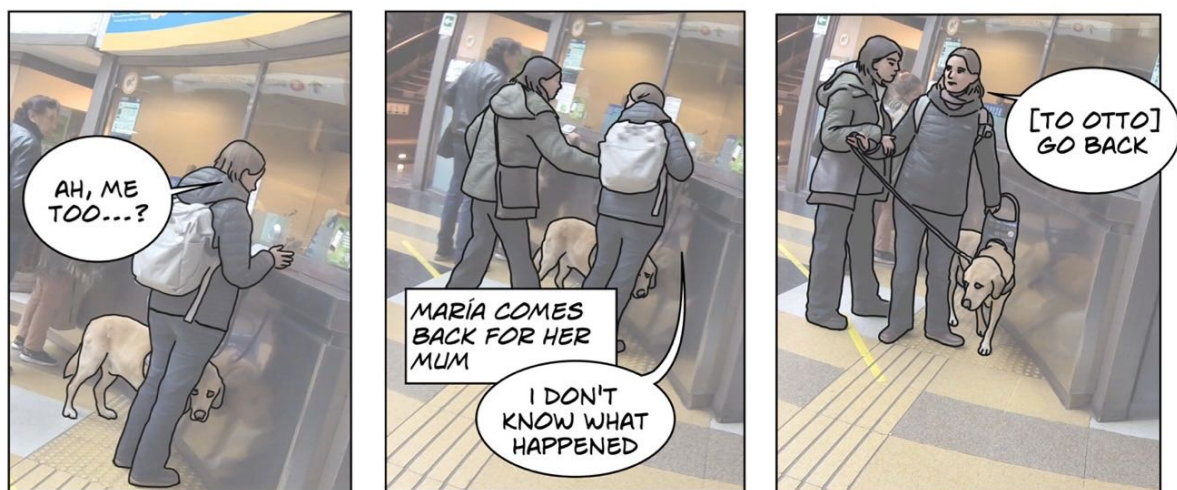


Figure 6: Tracing to anonymise and detail (source: Muñoz 2026)

Multiple cameras provide for multiple choices in transcript production. In the recording, which was the basis for figure 7, cameras were attached to the rucksack straps of walkers providing first person perspectives on the action. While these angles gives the reader access to the digital map the walkers were consulting, the person whose camera angle was used, was systematically missing from the visual record. In figure 7, the recording from the other camera was used as a screen grab in panel 3. For the reader it allows several things: the showing of the missing image of one of the party; a sense of the location of the two walkers in the landscape; and their relative position to one another. The latter is side-by-side when they compare their maps with one another. It's discernible in panel 2 and then made clear in panel 3. Ideally, we could have tracked their relative position throughout, it is only when then the walker with the hat, in panel 3, turns completely around which is when he provides a usable shot of the other two walkers looking at the map.



Figure 7: Screen grabs from two cameras (source: Laurier et al. 2019)

Returning briefly to how we can show timing in a graphic transcript. Figure 8, uses the length of the caption box to show a customer monitoring one staff member ('S2') drinking a glass of water. It shows how he maintains his orientation until the staff member finishes drinking before then making a joke that he's drinking gin rather than water.



Figure 8: Using a caption box to show timing (source: author)

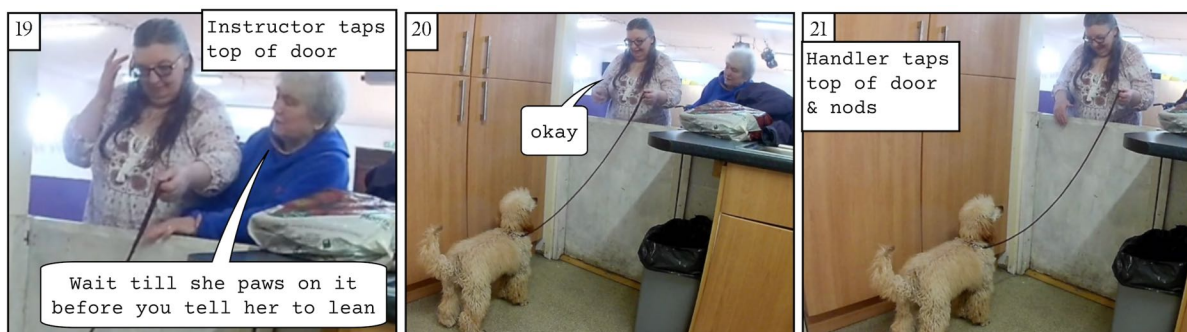


Figure 9: Using a crop to show a feature and to shift perspective (source: Laurier 2024)

A common tendency remarked upon by Murphy is to simply add the entire screen grab to transcripts without considering how it might be adjusted to better show the phenomenon. In graphic transcripts, the image is crop-able which thereby can provide either a close-up (figure 9, panel 19) or a more subtle shift of perspective (figure 9, panel 21) and again, one whose sense is assembled across the panels by the reader. In figure 9, panel 19 brings us a close-up of two dog trainers while we consider the instruction of one to the other. Panel 21 has the reader's attention directed to the interaction between one trainer and the dog. The reader is carrying over the details of the two trainers into and across panels 20 and 21 as the perspective widens out then narrows again to focus on the trainer and the dog.

What has hopefully become apparent across the examples used in this section is that there are multiple possibilities for making transcripts less ‘bad’. In gathering the collections of samples for this chapter it makes me realise that I have tended to use only one unfamiliar representational form at a time rather than multiple ‘experiments’ in transcription. One reason for doing so has been to retain the familiarity of the comic strip’s conventions. As Tyler (2021: 398) puts it: ‘to retain the integrity of the comic strip genre, the researcher is limited in the number of linguistic and prosodic details she can include’. A second reason is only bringing one feature into focus at one time rather than attempting to draw the reader’s attention to multiple interests in one transcript.

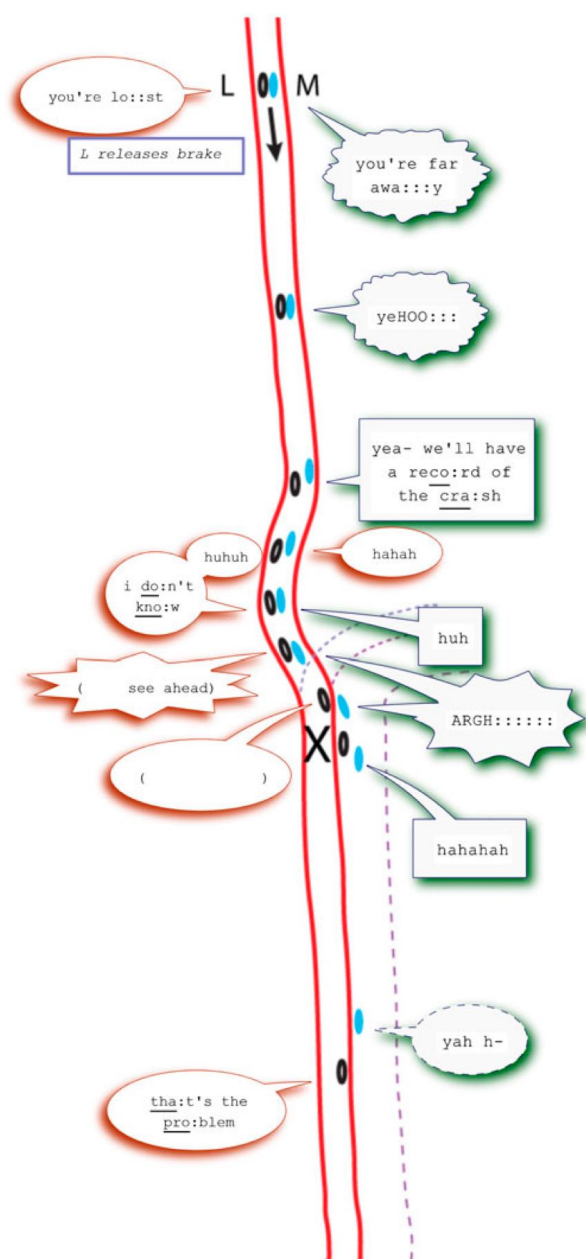


Figure 10: Transcribing location and talk (source: McIlvenny 2013)

Indifference, invitations and revelations

In reflecting on the conventions for textual transcripts, Mlynář and Rieser (2025) alert us to how transcribing has an abiding tension within it. It serves to make the familiar strange by the very practices of re-watching which are required and are, anyway, part of pursuing what people are doing via a recording of them doing it. Transcribing is seeking to recover the practices which people are using to produce, organise and make accountable what they are up to. The tension arises in the quality, which I have touched upon from the outset, of transcribing as making all kinds of things familiar through transcribing in, for example, Jeffersonian style. A process which allows other analysts to re-examine what happened, drawing upon the transcript as a good enough record of those earlier events.

As a final and very striking example Paul McIlvenny (figure 10) experimented with blending graphic transcripts and cartographic conventions (see also Madden 2005:94, for a more cartographic and humorous version). Studies of mobility like McIlvenny's have recurrently become testing grounds for graphic transcripts (and textual transcripts) because motion itself does not only have thing in motion it puts the location of the speaker as a consideration in understanding their actions. Just where the cyclists are and where they are going toward brings with them, a locationally-generated set of sense-making resources for speakers. Alongside timing for turn-taking it is examining and shows its examination of where, and towards-where, someone is saying something. How a twist in a road easily visible in figure 10 becomes a feature which is anticipated by the speakers as they approach it and them dramatized by them as they cycle through it.

Mlynář & Rieser argue that transcription should follow a spirit of wait and see, not weighed down by existing social science theories or techniques in 'which transcription is conceptualised as an invitation for both the audience and the researcher to exercise indifference to plausibilities and professional sensibilities that exist prior to seeing the data and to look instead at its possibilities, as wild as they may be' (Mlynář & Rieser 2025: 18). They direct us toward a willingness to wait and see not just what, but also *how* we will transcribe the phenomena that interest us. The invitation is not to be motivated by what seems tenable or probable, particularly in the light of existing disciplinary concerns. They are asking us as transcribers to value what the transcript reveals rather than judge it by its adherence to existing rules in one format or another. Their worry reveals as much about the dominance of certain formats of transcription which have come to be assumed as the markers of competent analyst rather than the pursuit of insights and revelations of phenomena.

References

- Ayaß, R. (2015). Doing data: The status of transcripts in Conversation Analysis. *Discourse Studies*, 17(5), 505–528. <https://doi.org/10.1177/1461445615590717>
- Bogen, D. (1999). *Order Without Rules, Critical Theory and the Logic of Conversation*. SUNY Press.
- Buscher, M. (2005). Social Life Under the Microscope. *Sociological Research Online*, 10(1), 100–123. <https://doi.org/doi.org/10.5153/sro.966>
- Harness Goodwin, M., & Cekaite, A. (2025). The Development of Video Analysis. In *The Routledge International Handbook of Ethnomethodology* (pp. 237–247). Routledge. <https://doi.org/10.4324/9780429323904-23>
- Hepburn, A., & Bolden, G. B. (2017). *Transcribing for Social Research*. Sage.
- Jefferson, G. (1978). What's in a 'Nyem. *Sociology*, 12(1), 135–139. <https://doi.org/10.1177/003803857801200109>
- Laurier, E. (2014). The Graphic Transcript: Poaching Comic Book Grammar for Inscripting the Visual, Spatial and Temporal Aspects of Action. *Geography Compass*, 8(4), 235–248. <https://doi.org/10.1111/gec3.12123>
- Laurier, E. (2019). The panel show: further experiments with graphic transcripts and vignettes. *Social Interaction. Video-Based Studies of Human Sociality*, 2. <https://10.7146/si.v2i1.11396>
- Laurier, E., & Arathoon, J. (2024). Understanding in the canine classroom: learning to 'lean'. *Scottish Geographical Journal*. <https://doi.org/10.1080/14702541.2024.2379020>
- Laurier, E., Brown, B., & McGregor, M. (2016). Mediated Pedestrian Mobility: Walking and the Map App. *Mobilities*, 11(1), 117–134. <https://doi.org/10.1080/17450101.2015.1099900>
- Laurier, E., & Boelt Back, T. (2023). Recurrent Problems and Recent Experiments in Transcribing Video, Live Transcribing in Data Sessions and Depicting Perspective. In P. Haddington, T. Eilittä, A. Kamunen, L. Kohonen-Aho, T. Oittinen, I. Rautiainen, & A. Vatanen (Eds.), *Ethnomethodological Conversation Analysis in Motion. Emerging Methods and New Technologies* (pp. 245–263). Routledge. <https://doi.org/10.4324/9781003424888>
- Laurier, E., Muñoz, D., & Swanton, D. (2024). Encounters convening publics: opposing politics and confronting trouble-making. *Social & Cultural Geography*, 25(10), 1670–1690. <https://doi.org/10.1080/14649365.2024.2357686>
- Smith, T. A., Laurier, E., Reeves, S., & Dunkley, R. A. (2019). “Off the beaten map”: Navigating with digital maps on moorland. *Transactions of the Institute of British Geographers*, 45, 223–240. <https://doi.org/10.1111/tran.12336>
- Ivarsson, J. (2010). Developing the construction sight: Architectural education and technological change. *Visual Communication*, 9(2), 171–191. <http://vcj.sagepub.com/content/9/2/171.short>
- Lindwall, O., & Lymer, G. (2024). Detail, Granularity, and Laic Analysis in Instructional Demonstrations. In M. Lynch & O. Lindwall (Eds.), *Instructed and Instructive Actions. The Situated Production, Reproduction, and Subversion of Social Order* (pp. 37–54). Routledge. <https://doi.org/10.4324/9781003279235>
- Luff, P., & Heath, C. (2015). Transcribing Embodied Action The Handbook of Discourse Analysis. In D. Tannen, H. E. Hamilton, & D. Schiffrin (Eds.), (pp. 397–321). Wiley Blackwell.
- Madden, M. (2005). *99 ways to tell a story: exercises in style*. Chamberlain Bros.
- McCloud, S. (1993). *Understanding Comics, The Invisible Art*. Kitchen Sink.
- McIlvenny, P. (2013). The Joy of Biking Together: Sharing Everyday Experiences of Vélomobility. *Mobilities*, 10(1), 55–82. <https://doi.org/10.1080/17450101.2013.844950>
- Mlynář, J., & Rieser, Y. (2025). Transcribing encounters with the 'wild': The neglected case of the poetics of ordinary talk. *Qualitative Research*, 0. <https://doi.org/10.1177/14687941251350878>
- Mondada, L. (2018). Multiple Temporalities of Language and Body in Interaction: Challenges for Transcribing Multimodality. *Research on Language and Social Interaction*, 51(1), 85–106. <https://doi.org/10.1080/08351813.2018.1413878>

- Muñoz, D. (2026) An infrastructure of embodied practices: How disabled people become part of public transport in Santiago de Chile. *Urban Studies*, online first, <https://doi.org/10.1177/00420980251413498>
- Murphy, K. M. (2021). Transcription Aesthetics. *Semiotic Review*, 9, 1–32. <https://doi.org/10.1515/cogsem-2019-2006>
- Pelikan, H. (2023a). Transcribing human–robot interaction. In *Ethnomethodological Conversation Analysis in Motion* (pp. 42–62). Routledge. <https://doi.org/10.4324/9781003424888-4>
- Pelikan, H., & Hofstetter, E. (2023b). Managing Delays in Human-Robot Interaction. *ACM Transactions on Computer-Human Interaction*, 30(4), 1–42. <https://doi.org/10.1145/3569890>
- Schegloff, E. (2000). On Granularity. *Annual Review of Sociology*, 26, 715–720.
- Smith, R. J. (2024). Fieldwork, participation, and unique-adequacy-in-action. *Qualitative Research*, 24(1), 60–80. <https://doi.org/10.1177/14687941221132955>
- Skedsmo, K. (2021). How to Use Comic-Strip Graphics to Represent Signed Conversation. *Research on Language and Social Interaction*, 54, 1–20. <https://doi.org/10.1080/08351813.2021.1936801>
- Stukenbrock, A. (2009). Herausforderungen der multimodalen Transkription: Methodische und theoretische Überlegungen aus der wissenschaftlichen Praxis. In K. Birkner & A. Stukenbrock (Eds.), *Die Arbeit mit Transkripten in Fortbildung, Lehre und Forschung* (pp. 144–169). Verlag für Gesprächsforschung.
- Tyler, R. (2021). Transcribing whole-body sense-making by non-dominant students in multilingual classrooms. *Classroom Discourse*, 12(4), 386–402. <https://doi.org/10.1080/19463014.2021.1896563>
- Wiggins, S., Majlesi, A. R., Ekström, A., Hydén, L.-C., & Cekaite, A. (2024). How assisted eating becomes a caring practice in institutional settings: Embodied gestures and stages of assisted eating. *Appetite*, 200, 107552. <https://doi.org/10.1016/j.appet.2024.107552>

Further Reading:

- Laurier, E. (2014). The Graphic Transcript: Poaching Comic Book Grammar for Inscribing the Visual, Spatial and Temporal Aspects of Action. *Geography Compass*, 8(4), 235–248. <https://doi.org/10.1111/gec3.12123>
- Laurier, E. (2019). The panel show: further experiments with graphic transcripts and vignettes. *Social Interaction. Video-Based Studies of Human Sociality*. 2. <https://doi.org/10.7146/si.v2i1.11396>
- Laurier, E., & Boelt Back, T. (2023). Recurrent Problems and Recent Experiments in Transcribing Video, Live Transcribing in Data Sessions and Depicting Perspective. In P. Haddington, T. Eilittä, A. Kamunen, L. Kohonen-Aho, T. Oittinen, I. Rautiainen, & A. Vatanen (Eds.), *Ethnomethodological Conversation Analysis in Motion. Emerging Methods and New Technologies* (pp. 245–263). Routledge. <https://doi.org/10.4324/9781003424888>
- Murphy, K. M. (2021). Transcription Aesthetics. *Semiotic Review*, 9, 1–32. <https://doi.org/10.1515/cogsem-2019-2006>