

# Reflections on transcription<sup>1</sup>

by Paul ten Have, University of Amsterdam

In this essay, I will present some reflections on transcription, as it is practised in conversation analysis (CA). In my opinion, transcription is an essential part of the CA research process, but it is, at the same time, deeply problematic. In order to understand both its essential character and its problematic features, we should first consider the particularities of CA as a research enterprise.

## CA and recordings

Conversation analysis, in the sense used here, emerged in the mid-1960s as some young sociologists working in California struggled to find a way of ‘doing sociology’ that would fit their particular interests and preferences, while avoiding some of the difficulties the then current approaches were displaying in their opinion. Both Harvey Sacks and Emanuel Schegloff had been working with Erving Goffman as part of their graduate studies at the University of California at Berkeley, while Sacks had developed a working relationship with Harold Garfinkel at the University of California in Los Angeles. The insight that many issues that had been debated for centuries in the human sciences could be elucidated by paying close attention to the details of human interaction in situations of *co-presence* clearly reflected Goffman’s influence. But their approach to interaction differed from his as they were much more critical of the then current intellectual styles in sociology than Goffman ever was. In that respect, Harold Garfinkel’s efforts to build an ‘alternate sociology’ under the name of ‘ethnomethodology’ had a major impact on the intellectual particularities of what later became conversation analysis. To summarize in a few words, Garfinkel objected to the practice in sociology of studying the order in and of society through the conceptual ordering which sociologists had themselves applied, while ignoring this constitutive effect, as well as the constitutive activities of ordinary societal members. Therefore, the task of ethnomethodology was to study ‘members’ methods’, the everyday, ‘seen but unnoticed’ ways in which members of society, social scientists included, constituted the facts of society as part of their ordinary activities. In a sense, then, Sacks and Schegloff used Garfinkel’s program of an alternate sociology to study the kinds of phenomena that Goffman had put on the sociological agenda. In so doing, however, they developed a style of social research that was remarkably different from both sources, Goffman and Garfinkel.<sup>2</sup>

An essential part of the ethnomethodological program is the effort to study members’ practices as such, rather than some sociological work-up of their products, as in survey research tables or archived documents. This is not an easy matter, as the constitutive aspects of those practices are,

---

<sup>1</sup>Text based on a talk, given at the Colloque “l’Oral en contexte: des objectifs aux méthodes”, Université Paul Valéry - Montpellier III, 17 November 2001.

<sup>2</sup>The text of this paragraph sums up a rather complicated set of developments and relationships that would require a much more extensive treatment than can be given here; earlier and fuller discussions are available elsewhere (Ten Have, 1990, 1999). For Schegloff’s treatment of the development of CA, focussing on Sacks, see his introductions to the Sacks lectures (Sacks, 1992).

for members (again including sociologists) ‘essentially uninteresting’, and therefore hard to get in focus. Working with tapes and transcripts can be seen, then, as not just a practical way of getting detailed data, but also as a solution to the problem of ‘the invisibility of common sense procedures’ (Ten Have, 1990, see also Clayman & Maynard, 1995).

A couple of quotes from Harvey Sacks’ lectures may help to further clarify the intimate relationship between CA’s purposes and its methodological practices.

When I started to do research in sociology I figured that sociology could not be an actual science unless it was able to handle the details of actual events, handle them formally, and in the first instance be informative about them in the direct ways in which primitive sciences tend to be informative - that is, that anyone else can go and see whether what was said is so. And that is a tremendous control on seeing whether one is learning anything. [...]

I started to work with tape-recorded conversations. Such materials had a single virtue, that I could replay them. I could transcribe them somewhat and study them extendedly - however long it might take. The tape-recorded materials constituted a "good enough" record of what happened. Other things, to be sure, happened, but at least what was on the tape had happened. It was not from any large interest in language or from some theoretical formulation of what should be studied that I started with tape-recorded conversations, but simply because I could get my hands on it and I could study it again and again, and also, consequentially, because others could look at what I had studied and make of it what they could, if, for example, they wanted to be able to disagree with me.

Sacks (1984a): 26; from a lecture given in the fall of 1967.

So, for Sacks, working with tape-recorded conversations had a kind of exemplary value of making the details of actual human action available for detailed scrutiny and formal analysis. For him, that meant being able to formulate ‘rules’ that would ‘provide for’ the observed details, and ‘yield the technology of conversation’:

The idea is to take singular sequences of conversation and tear them apart in such a way as to find rules, techniques, procedures, methods, maxims ... that can be used to generate the orderly features we find in the conversations we examine. The point is, then, to come back to the singular things we observe in a singular sequence, with some rules that handle those singular features, and also, necessarily, handle lots of other events.

Sacks (1984 b): 413; from one or more lectures given in 1970.

So the main function of recordings is to provide access to the details of human conduct in general, and interaction in particular, in the first instance for the researcher, and secondly also to his or her audience. Recordings of human interaction are, in CA, not just listened to or looked at, they are also rendered in textual formats as ‘transcription’, and it is reflection on this latter process that is the main task of this essay.

Transcription as part of the CA research process

The activity of transcription constitutes a particular phase in the process of doing conversation analysis, as depicted in the following sequential schema<sup>3</sup>:

Original (inter-)action → *recording* → (audio/video-)record → *transcription* → transcript → (action) *understanding* → procedural *analysis* → analytical argument

In this schema, the *italicised processes* are selectively reductive vis à vis the preceding states/products. My interest in this essay is to consider the specific properties of these selective reductions, which can be seen as ‘losing’ features of the preceding state and/or as focussing on (and foregrounding) features of specific interest. When ‘looking forward’ the processes may be seen as instrumental in gaining a sharper focus on the phenomena of interest, which were already present in the preceding state. When you ‘look backwards’, however, you will have to admit that you cannot reconstitute the earlier state from the later rendering, because features that may have been essential in constituting the earlier state in its full richness are no longer available in the later rendering. This is another version of the asymmetrical properties of the action-account pair, as often noted by Harold Garfinkel (cf. Garfinkel & Wieder, 1992).

The purpose of the first two processes, i.e. *recording* and *transcription*, is to produce a non-perishable, transportable, and manageable representation – an ‘immutable mobile’, in Bruno Latour’s terms (1987: 228) – to assist in the later following processes of *understanding* and *analysis*.

Gail Jefferson, who designed the CA transcription conventions, starts her 1985 essay on the transcription and analysis of laughter as follows:

I take it that when we talk about transcription we are talking about one way to pay attention to recordings of actually occurring events. While those of us who spend a lot of time making transcripts may be doing our best to get it right, what that might mean is utterly obscure and unstable. It depends a great deal on what we are paying attention to. It seems to me, then, that the issue is not transcription per se, but what it is we might want to transcribe, that is, attend to.

Jefferson (1985): 25

In other words, the inevitable reduction, simplification and idealization which are the effect of these processes, have to be considered in terms of the specific analytic interests that are brought to bear on the original. In other words, before discussing the cost of the inevitable losses which the two processes of *recording* and *transcription* bring about, one has to clarify which aspects, properties or features of the original will have to be analyzed and explicated. In short, one has to be clear about one’s analytic object<sup>4</sup>.

---

<sup>3</sup>This schema was partly inspired by Ashmore & Reed (2000), which will be discussed in a later part of this paper.

<sup>4</sup>For reasons of space and economy, I will not discuss the specific ‘losses’ that occur as part of the recording process.

## On description versus transcription

In her 1985 essay on the transcription and analysis of laughter, Jefferson contrasts, referring to every day occasions, the treatment *in* subsequent talk *of* previous talk and of previous laughter: while talk may be quoted (and perhaps even mimicked) laughter does not seem to be ‘quotable’ to the same extent. Similarly, in transcripts, laughter used to be *described* rather than *transcribed*. Let me start my discussion of this contrast with a digression.

## Bird song depictions in field guides

Consider what the writers of field guides for bird watchers do when they discuss bird songs as a property of a species.

From: P.T. Peterson et al, *Petersons Vogelgids van alle Europese vogels*.

### KLEINE KAREKIET [Reed Warbler]

...

Geluid: een laag *tsjur*, een scherp, alarmerend *skurr* (als dat van Rietzanger) en een zwak tikkend geluid. Aangehouden zang lijkt op die van Rietzanger, maar is meer herhalend en maatvast: *tsjirruk-tsjirruk, djek, djek, tirri-tirri-tirri*, vermengd met vloeiende en nabootsende geluiden. Zingt overdag en ‘snachts.

[Voice: a low *tsjur*, a sharp, alarming *skurr* (as that of Sedge Warbler) and a weak ticking sound. Prolonged song similar to that of Sedge Warbler, but is more repetitive and steady: *tsjirruk-tsjirruk, djek, djek, tirri-tirri-tirri*, mixed with flowing and imitating sounds. Sings in daytime and at night.]

From: Kilian Mullarney et al, *ANWB Vogelgids van Europa*

### Kleine Karekiet

GELUID Roep een kort, onopvallend *tsje*, soms iets harder, bijna smakkend *tsjk*. Bij opwinding een langgerekt, schor *sjrieh*, een vet, rollend *sjrrre* en een tweelettergrepig *trr-rr*. Zang 'babbelend' in laag tempo, bestaand uit nerveuze, 2-4 keer herhaalde noten (onomatopoëtisch), af en toe onderbroken door imitaties of fluittonen, *trett trett trett TIRri TIRri truu truu Tle tre tre wi-wuu-wu tre tre truu truu TIRri TIRri ....* Tempo af en toe hoger, maar nooit met crescendo van Rietzanger.

[SOUND Call a short, unremarkable *tsje*, at times a bit louder, almost smacking *tsjk*. In excitement a long drawn, hoarse *sjrieh*, a fat, rolling *sjrrre* and a two-syllable *trr-rr*. Song ‘babbling’ at a slow tempo, consisting of nervous, 2-4 time repeated notes (onomatopoetic), now and then interrupted by imitations or whistlings, *trett trett trett TIRri TIRri truu truu Tle tre tre wi-wuu-wu tre tre truu truu TIRri TIRri ....* Tempo now and then higher, but never in crescendo like Sedge Warbler.]

Note in these examples a mixture of *descriptions* and some efforts at *transcription*, with for the same species rather different results! The purpose of the *transcriptions* is, of course to compensate for the limited success of *descriptions* for the purpose at hand, making actually heard calls and songs identifiable as produced by specific species of birds. The language of humans is of limited use in providing a recognizable image of calls and songs produced by birds. In the same vein, laughter done by humans seems to be difficult to ‘picture’ as well, as we will see next.

Returning to Jefferson’s essay, I quote two different versions of transcripts by her of the same recording:

(7) (GTS:1:1:14, 1965)

Ken: And he came home and decided he was gonna play with  
his orchids from then on in.

Roger: With his what?

Louise: heh heh heh beh

Ken: With his orchids. [He has an orchid-

Roger: [Oh heh hehheh

Louise: ((through bubbling laughter)) Playing with his organ yeah  
I thought the same thing!

Ken: No he's got a great big [glass house-

Roger: [I can see him playing with his  
organ hehh hhhh

Jefferson (1985): 28

(GTS:1:2:33:r2, 1977)

Ken: An'e came home'n decided'e wz gonna play with iz o:rchids.  
from then on i:n.

Roger: With iz what?

Louise: mh hih hih [huh

Ken: [With iz orchids.=

Ken: Ee[z got an orch[id-

Roger: [Oh:. [hehh[h a h 'he:h] 'heh

Louise: [heh huh 'hh] PLAYN(h)W(h)IZ 0(h)R'N

ya:h I [thought the [same

Roger: [uh::: ['hunhh 'hh 'hh

Ken: [Cz eez gotta great big[gla:ss house]=

Roger: [I c'n s(h)ee

Ken: =[(

Roger: =[ im pl(h)ay with iz o(h)r(h)g'(h)n 'uh

Jefferson (1985): 29

The crux of Jefferson’s argument is that the later transcription allows one to analyze the interaction taking place in greater depth, because it provides details of timing and inter-action that are not available in the first rendering. In the case at hand, she suggests, it does not seem to be an accident that the girl laughs through the obscenity, producing it in a suggestive but not well-

articulated manner, while continuing afterwards in an undisturbed voice. Extending her argument, one can suggest that the standard orthography rendering of spoken interaction, i.e. in the language of writing, is a poor means to picture the hearably functioning details of that interaction.

### Illustration

When CA researchers start working on a transcription task, they are faced with a number of dilemma's. Any actually produced transcription is analyzable as a practical but always ambivalent solution to inescapable dilemma's in transcription routines.

- The use of standard orthography, with more or less adaptations to display some of the properties of the actual speech production: 'words-as-spoken' versus 'sounds-as-uttered'
- The use of mechanical timing devices for pauses, versus a reliance on informal procedures like counting syllables in muttered words, as an unavoidable subjective 'measure' that may take into account pace relativity
- Decisions regarding formatting issues, for example line breaks to signal 'describable actions' versus a more continuous rendering.

I have discusses these and other practical issues of doing transcriptions elsewhere (Ten Have, 1999: 75-97; cf. also Psathas & Anderson, 1990). For now, I will just illustrate some of the issues raised so far on the basis of an extract from my own research. It has been taken from a transcript of a recording of a medical consultation made in the Netherlands in the late 1970s. A mother consults with her daughter. She has described her daughters complaints in lay terms and then the physician has asked the girl to show him her tongue. After some more descriptions from the mother and one question/answer exchange with the daughter, the physician provides a preliminary diagnosis as follows:

### Extract 1

- 54 A: 'hh nou we zullen es kijken,  
 54 A: 'hh well we will take a look  
 55 A: d't kan eh (0.5) †eenvoudig (0.9) 'te zijn=  
 55 A: it can uh (0.5) simply (0.9) be  
 56 A: =>dat ze (bevoorbeeld) wat tekort aan bloed heeft.<  
 56 A: that she has for instance a little blood shortage  
 57 A: ze is [†negen †jaar,  
 57 A: she is [nine years  
 58 M: [(°ja heb ik ook al°)  
 58 M: [(°yes I have also already°)  
 59 A: 'hhh de leeftijden †één jaar †vier jaar †negen,  
 59 A: 'hhh the ages one year four years and nine  
 60 A: ja tien elf >zo'n beetje rond-tie tijd,=  
 60 A: yes ten eleven araound that time  
 61 A: =als ze een beetje< †uit gaan schieten.  
 61 A: when they begin to grow  
 62 A: 'hhh dat zijn >tijden waarop kinderen vaak=  
 62 A: 'hhh those are times when children often

- 63 A: =een beetje ↑bloedarmoede [hebben.  
 63 A: *have a little blood sh[ortage*  
 64 M: [(jjh) twee jaar >ge↑leden=  
 64 M: [(jjh) two years back  
 65 M: =heeft ze 't ↑ook gehad,=  
 65 M: *she also had that*  
 66 M: =toen ↑ook in september,=  
 66 M: *also in September then*  
 67 M: =toen waren we bij de ↑schoolarts,=  
 67 M: *when we visited the school doctor*  
 68 M: =en toen had ze ↑ook bloed[armoede.<  
 68 M: *and she also had a blood [shortage then*  
 69 A: [↓hmm  
 69 A: [↓hmm  
 70 (1.6)  
 71 A: >'k wee- niet of het wat ↑is=  
 71 A: *I don't know whether it's something*  
 72 A: =maar we kunnen ('t) even (↓prikken).<  
 72 A: *but we can just prick*  
 73 (1.4)

As a reader of this transcript, you take on a kind of virtual overhearer's perspective. What you see is a rendering of speaker A talking in lines 54-57, then a short and incomplete contribution by speaker M, partly overlapping the A's talk (58), A's continuation in lines 59-63, and just before he is finished, M taking up again, continuing for a few lines (64-68), and just before she finishes, a short 'hmm' from A, then a pause, and finally A starting to talk again (71 and following).

Using the contextual information I provided, you know that A is a physician and M the mother of a young patient. From line 54 onwards, the physician 'has the floor', which he uses to announce a further action (54) and a preliminary diagnosis (56). In line 58 the mother mutters something which I have rendered as '°ja heb ik ook al°', and translated as '°yes I have also already°'. This utterance is obviously not complete, but it can be plausibly expanded into 'ja heb ik ook gedacht', *yes I've been think of that also already*. The doctor does not hearably/visibly react to this muttering; he may not have heard it or he may have chosen to ignore it. In any case, he continues his explanation (57, 59-63), suggesting that the diagnosis may fit in an age-related pattern. Something similar to the earlier muttering happens in line 64, but this time the mother gets the floor, to refer to an earlier experience with a equivalent complaint, which was diagnosed by another doctor. The physician reacts to this in a minimal fashion '↓hmm' (96), then there's a pause, after which he initiates a new phase in the encounter (71).

In such an overall hearer's/reader's description, it is hard to avoid action ascriptions. The overall theme in the just given description is one of turn-taking (Sacks et al, 1978). And it is in terms of turn-taking that most of the CA-specific details in the transcription gain their significance. It is in these terms that one can speak of having the floor, producing a secondary speaker remark, keeping the floor, changing speakership, etc. The turn that A takes in lines 55-56 can be heard as complete, both in terms of propositional content, and of intonation: line 55 is produced hesitantly,

while 56 is faster and it ends with a downward, final intonation. Therefore, the mother may have taken his announcement as finished, although in fact it isn't. As she starts her comment a bit slow, the physician can continue talking. She solves the overlap problem by turning silent before she is finished, although she was able to produce a word or two in the clear. The physician seems to 'accept' her overlap solution by producing a hearable inbreath before he continues his explanation. And again, the transcriptional details provide us with the materials to understand the next speakership change in lines 63/64. We can analyze the explanation's semantic structure to propose that it is possibly complete at that point, while the intonation contour, with a stress on the pre-final key term, and the downward ending of the last one 'confirms' such an analysis. The mother, however, does not even wait for this final word and produces a semantically empty pre-start item, before she makes another remark, relatively fast and without pauses.

What I have just given can be characterized as a technically informed effort at an 'action understanding' of this small episode. It is technically informed in that I use the CA transcription conventions to point at particular kinds of production details which 'invite' an understanding of the interaction in terms of turn-taking or 'floor management'. What I have done, then, is to use some theoretical and methodological 'tricks of the trade' of CA to elucidate the episode as a negotiation of turns-at-talk. We see the physician keeping the floor for some time and the mother 'watching' him, looking for a useable opening to insert her comments in. Further analyses, for example using ideas from that other Sacksian tradition of Membership Categorization Analysis (MCA – Sacks, 1972 a&b; Hester & Eglin, 1997), could be added to it. In MCA terms we can say that we see/hear the physician announcing a diagnosis and the mother inserting her comments of recognition of it as 'another case of what I thought it would be'; that is, the physician is doing his category-bound job, while the mother offers a display of her lay understandings.

Whether we use CA or MCA or both, we start with an overhearers' perspective and then try to use the information we are able to get to reconstruct the participants' perspectives as enacted in the 'overheard' interaction. What we as analysts do, then, is trying to convince our readers of the plausibility of this action understanding and the analysis that is based on it, referring to the utterances' properties *foregrounded* by our transcript's details.

In my exemplary analysis, above, I have not used all of the transcript details. In other words, my analysis has not 'exhausted' my transcript. For instance, the intonational information given might be used as grounds for a further analysis of the internal organization of the various turns-at-talk. In the extract's first turn, the part given in line 54 is produced in ordinary pace, the next one in line 55 is 'slower' or 'hesitant' with an 'uh' and two small turn-internal pauses, while the last part on line 56 is 'latched' to the previous part and produced more quickly. One might suggest that there is a certain parallelism here between these production details and the semantic message of these three parts: the first an unproblematic announcement of an upcoming examination, the second an indication of the hypothetical quality of the diagnosis, and the third the actual 'possible diagnosis', with an inserted 'for instance' and the quick pace stressing its 'dismissable' character. Next, the low volume of the mother's unfinished inserted remark may be related to both its quality as an insertion in overlap in the physician's turn, and its semi-private, 'lay' character. When we look at the rhythm of the next two utterances, first by the physician (59-63) and second by the mother (64-8), we can see how they stress the essential and/or enumerative elements in their contributions. The mother's turn, for instance, can be analysed as a three-part list (Jefferson, 1990), with the core elements in lines 65, 66 and 68, and an explanatory insertion in line 67. This structure is



punctuated, so to speak by the three times stressed *also*'s in the core parts. Together with their latched production and continuous intonation, these features make this into a 'strong', hard to interrupt (or ignore) package.

The analytic suggestions given in this section could be elaborated further in various ways. One could discuss comparative instances to substantiate the various claims as to the functional significance of the features discussed. Or one could use these observations as contributions to an analysis of the local accomplishment of, or negotiations about, institutional relationships (Ten Have, 2001). In the present context, however, the purpose was to offer a restricted demonstration of the analytic fruitfulness of using the Jefferson conventions as a kind of perceptual and thereby analytic shopping list.

### Critiques

In this section, I will discuss some issues that have been raised in various critiques of CA practices<sup>5</sup>.

A rather general ethnomethodological critique of CA's transcription practices can be found in a chapter called 'The organization of talk' in a book by David Bogen (1999)<sup>6</sup>. The book as a whole contains a sustained critique of Habermas' critical theory from a stance based on ethnomethodology and Wittgenstein's later philosophy. In this particular chapter Bogen develops a similarly inspired critique of the scientific features of conversation analysis<sup>7</sup>. The crux of this critique is that CA presents itself as a 'foundational' science, i.e. a science which abstracts generic and basic features of phenomena from accidental, singular and unique characteristics. Here is his core remark on transcriptions:

Transcripts are a pervasive and elementary feature of conversation analytic practice. One learns that practice in and through learning to transcribe. Whenever findings are presented, analysts proffer transcripts and then instruct readers in methodic, often ingenious ways of reading them as evidentiary support for their arguments. It is in this sense that transcription comprises a primitive literary technology of conversation analysis. Clearly the transcript contributes *something* to the arguments and demonstrations of conversation analysts, though what *specifically* it contributes remains unclear. The suggestion here will be that the transcripts function as a literary genre, the business of which is to establish the *actuality* of the events in question. What the transcript contributes to analysis is not the real events, but the literary analogue of the real events-what real events *are*, or *must be*. That is, the transcript lends to the analysis what Barthes has termed a

---

<sup>5</sup>I will focus, in this section, on only a few of the more relevant and general critiques of CA transcriptions practices.

<sup>6</sup>An earlier version of the chapter was published as Bogen (1992).

<sup>7</sup>As such, Bogen's chapter is a member of a family of ethnomethodologically inspired critiques of CA's practices and rhetorics, which includes Lynch (1993, esp. pp. 203-64) and Lynch & Bogen (1994; 1996).

"realistic effect."

Bogen (1999): 90

Referring to Gérard Genette, he compares the CA transcript to a Homeric style of narrative, which, 'also' has a mass off excessive details and redundant information. As Genette has it, such detail 'serves no purpose other than to let us understand that the narrative mentions it only because *it is there*, and because the narrator, abdicating his function of choosing and directing the narrative, allows himself to be governed by "reality," by the presence of what is there and what demands to be "shown" (cited in Bogen, 1999: 91). In a similar way, the level of detail present in a CA transcript would be, according to Bogen always be 'unavoidably *excessive* with respect to any particular analytic point being made'. Furthermore, the 'reality' evoked by a transcript's detail is not the original event, but rather 'the electromagnetic record'. Without such a record, the details in the transcript could never be recovered. It is the existence of a record, then, that is demonstrated in the transcript.

From a CA point of view, it may be remarked that Bogen's complaint – evident as such in his choice of terms like 'excessive', 'surplus detail', or 'hyperabundance of detail' – glosses over the potential analytic sense of those details (as I have tried to demonstrate in the previous section). As he writes in the passage quoted above: 'Clearly the transcript contributes *something* to the arguments and demonstrations of conversation analysts, though what *specifically* it contributes remains unclear.' That this contribution remains unclear *to him* may be due to his lack of 'vulgar competence', to use a Garfinkel term<sup>8</sup>, in the art of conversation analysis. It may be suggested that his 'escape' to an external literary analysis may be seen in connection with such a lack of competence. It is, of course, quite often the case that a large part of the details given in a transcript included in a CA publication is not taken up in the analysis as represented in that publication. Some practitioners have suggested to make full transcripts during the exploratory stages of a project, and simplify these in specific publications, deleting details that have not been taken up. Others may argue, however, that by keeping such details, readers are given the means to do partly independent analyses of their own, and confront these with the ones offered by the author. The latter may, for instance, have given a particular analysis of the action performed by an utterance that has been produced in overlap with another. By noting the more or less exact point of overlap, a reader may find grounds to propose a different analysis, even when the point of overlap has not been discussed by the author.

As noted elsewhere, transcription has various functions within the CA enterprise: 'making transcriptions helps to take note of particular phenomena, it serves to built an accessible data archive, and it provides an audience with a limited but useful access to the phenomena discussed in an analysis' (Ten Have, 1999: 78). The first two functions are best fulfilled by making more or less 'full' transcripts, using the Jefferson conventions as a useful shopping list of phenomena that have been proved potentially significant within the CA tradition. As suggested, one may or may not decide to preserve this fullness in publications, but a critical analysis of the CA enterprise that focusses so much on the supposed hyperabundance of transcript details, misses some crucial points

---

<sup>8</sup>Having a 'vulgar competence' in a trade is an essential part of the 'unique adequacy requirement' which, according to Garfinkel, is essential to ethnomethodological studies of that particular trade (cf. Garfinkel & Wieder, 1992: 182-4).

and stays on the literary surface of its phenomena.

I will now turn to the most sustained, while still more or less ‘friendly’ critique of CA transcription that I know, a recent (2000) paper by Malcolm Ashmore and Darren Reed, called: ‘Innocence and Nostalgia in Conversation Analysis: The Dynamic Relations of Tape and Transcript.’ It involves an epistemological investigation into the ways in which the two core ‘analytic objects’ of CA, the Tape and the Transcript, function in the CA research process. They first note that there is hardly any attention in the CA literature for the process of (audio) taping:

A plausible reason for this neglect, is that audio recording (and its results) has become so culturally naturalised, that, like photography, it is extremely difficult to problematise, to loosen the hold of its stubborn realism” (Par. 7<sup>9</sup>).

The origin of the Tape—its relation to any particular Event—is not of specific interest. This lack of interest in the process of recording (whether as a technical or conceptual issue) is an important first step in the "naturalising" of the Tape. In effect it provides for the "forgetting" of the Event, and its wholesale replacement by the Tape. (Par. 9).

Ashmore and Reed contrast this ‘realist’ position of the Tape, to the ‘constructivist’ treatment that the Transcript receives in the CA literature:

the Tape can be used as the standard against which the Analysis can be checked; and it can be revisited to produce a new Analysis, i.e. the Tape can be the source of more than one series of analytic objects (Par.10)

While transcription

[...] is routinely understood as a craft process, as itself a part of the practice of analysis, as conventional and constructive. Debates about "how much" to transcribe [...], or about the consequential features of particular systems and designs [...] are commonplace. Students (and other researchers [...]) are regularly warned not to fetishize the transcript, nor to treat it as the data, [...]. (Par. 11)

After discussing some differently styled evaluations of transcriptions versus tapes, they conclude that:

... the value of the transcript makes itself felt most clearly in the business of building the series of analytic objects that make up the "material" of any CA research project and thus in the search for *analytic utility*. On the other hand, when the tape appears as the "better" object of the two, what is being alluded to is its value in strengthening the *evidential utility* of the already-produced objects. (Par. 15)

---

<sup>9</sup>As is usual for online publications, detailed references for quotes in this paper are given in terms of paragraph numbers as provided in the electronic version, rather than to pages as in references to printed publications.

In other words, transcripts tend to be used when CA researchers are in a phase of searching for interesting phenomena to analyze, while tapes are the more trusted objects as it comes to proving one's analyses<sup>10</sup>.

They then develop some general schemas that are too complicated to be explained here, which they use to explicate some of the tensions and dynamics of CA work. To summarise: they contrast a left pole of *life* as apprehended in what Schutz has called the natural attitude with a right pole of *work* which is done in what they call 'the mediated attitude', which roughly corresponds to the Schutziian 'scientific attitude'<sup>11</sup>. Going from the left to the right, from the original event, via the tape and the transcript to the analysis and finally the CA publication, *life* is more and more transformed into *work*. At any moment, however, one can refer back to a more leftward item, which Ashmore & Reed refer to a 'nostalgia dynamics'. In other words, a rightward move is a 'constructivist' one, while a leftward reference is by contrast 'realist'. In CA both are used in various ways, as in the 'mutual elaboration of Tape and Transcript'.

The various backward shifts [...] (from Article to Analysis, from Analysis to Tape, etc.) can be understood to be doubly motivated. First, there is the operation of the general nostalgia dynamic, prompting returns to earlier, and thus more "actual", more "lifelike", stages of the analytical process. What is sought for, in effect, is the recovery of some level of "original detail". Closely connected is the desire to revisit the past for purposes of strengthening the evidential adequacy of the analysis, by checking (say) the Analysis against the Transcript, or the Transcript against the Tape. In each backwards shift motivated in this way, the "earlier" object is treated as (relatively) fixed with respect to the "later". Indeed, on such occasions, the former acts as an unquestioned standard with which to assess the fidelity of its translation into the latter. On these occasions, then, the leftward analytic object is reified. (Par.27)

A further contrast is built between two kinds of epistemic apprehension of objects, characterized as a relatively spontaneous hearing or seeing, versus a more studied listening or reading. They apply this contrast to the process of transcription in the following way:

CA's rhetoric of method generally understands the relationship of Transcript to Tape in terms of "representation" [...] or translation from one modality (aural) to another (textual). In order to achieve this translation as "faithfully" as possible, the Tape undergoes intense and focussed *listening*. The interpretative and productive act of *listening* changes the Tape's status from an unknown to a known, from an

---

<sup>10</sup>In their paper, the authors quote me as preferring transcripts over tapes as useful objects. This position can be related to my stress on the heuristic aspects of doing CA, both in the passage which they quote and in the illustration of a transcript's usefulness given in the present paper. In terms of the authors, I have stressed a transcripts's 'analytic utility'.

<sup>11</sup>They write in their note 4: "We prefer 'mediated' as this term can account for all activities of formulation, understanding, representation, performance, whether done in the course of scientific work or not.

object that is radically unstable to one which is relatively fixed. *Listening* polices the Tape. The "rules for hearing" distilled from this process are articulated in and as the Transcript. Thus, at this stage, the Transcript appears not so much as the Tape's translation, but as its *caption*. In Bruno LATOUR's terms, the "coupled object" of Tape & Transcript, bound together as "image & caption", have begun to take on the character and utility of an "immutable mobile" (LATOUR 1987). As we will see, the immutable character of the Tape is strengthened in subsequent turns (Next Times) which construe it as *heard* rather than *listened* to. (Par.34)

In short, in and through the rightward leaning activity of attentive listening, the transcribed tape gets its overall character as a natural object, which can then subsequently be heard 'as it is', in a leftward move of 'nostalgia'. This has important consequences for CA's claim that 'anyone can see for themselves' whether an analysis is correct, by checking 'the data'. Against this the authors argue that once 'instructed' by the author's transcript and analysis, the readers cannot achieve a naïve observation similar to the original 'innocence' of the author at his or her First Time Through: 'They are simply not in a position to approach the task of reanalysis with the requisite innocence'. (Par. 44)

Furthermore,

... the reader of a "researcher's report" always has much less information than the author. When a written piece of analysis is presented to a wider audience, it is fragments of transcript that evidence the author's analytic claims. Should the reader wish to "go further" than these texts allow, s/he will have to "go back", on our nostalgic trajectory, to a more complete set of materials—the (whole) Transcript, the Tape—which is always somewhere else. It is this problem which motivates [some] analysts [...] to advocate a digital solution: a transcript-free hypertext linking the Analysis directly to the Tape. (Par. 45)

The crux of this paper, then, can be read as an invitation to CA researchers to critically consider and openly reflect on the mix of realism and constructivism in their established practices and rhetorics. I have personally no problem to see my work in CA as no more than an analytic work-up of some 'slices of life' (or 'spates of talk') which, by being taken out of their original context of a lived stream of co-experience and transformed into analytic objects, inevitably have lost their primary significance. In other words, as analysts we can only use our 'overhearer's perspective, to re-construct a plausible version of the conversationalists' actually lived participant's perspectives. Within that process, producing a transcript is an essential part of this 'instrumentalization', about halfway between the poles of 'life' and 'work'.

## Conclusion

Before I conclude this essay, I would like to return for a moment to my earlier digression concerning the description and transcription of bird songs and calls. The examples I quoted and discussed were taken from a particular pragmatic context: field guides to be used by lay or

professional ornithologists as an aid in the identification of species of birds<sup>12</sup>. Such usage is based on the assumption of identifiable ‘species’, i.e. sets of birds that are willing and able to mate and produce fertile offspring. Species, then, are the theoretical objects on which the usage of field guides is oriented. The pragmatics of bird species identification by songs and calls abstracts from individual and local, or as one might say ‘cultural’ intra-species variations in order to focus on the differential identification of the species. ‘A species’ is always and inevitable a momentarily ‘fixed’ construction, a ‘violent’ cutting-up of the immense variability of life. The proliferation of ‘sub-species’ in recent field guides as well as phenomena of bastardation attests to the relative arbitrariness of species distinctions.

This analysis can be used to refocus on the pragmatic context of CA transcription. It may be suggested that the theoretical object which is the target of a CA transcription is the set of core devices that has been so far identified in the corpus of CA inquiries. The Jeffersonian transcript conventions represent the accumulated wisdom of the first generation of CA researchers as to the kinds of phenomena that would be good candidates for a CA type of analysis. Individual and local specifics of the recorded sound production are, of course, to be noted to a much greater extent than is done in field guides. But still, making the core phenomena of CA interests – such as the organization of turn-taking, sequencing, repair, etc. – is still a major function of a CA transcription’s selectivity.

However, although this overall orientation to CA’s core phenomena seems to be the guiding principle of CA transcription work, two related but distinct abilities are required to bring off useful transcriptions. These are the ability:

- to recognize words
- to clearly hear sounds.

The first requires knowledge of a language’s vocabulary embedded in the ability to understand spoken language in terms of its written analog. In that sense transcription is really ‘textualization’: translating oral language into written language. This phase of hearing what was said involves a kind of applied members’ work, in which the transcriptionist relies on his or her ‘ordinary’ or ‘vulgar’ competence as a member of a particular linguistic community.

The second requires the ability to distance oneself to a certain extent from the ‘official’ language, to hear the sounds as actually spoken. This would seem to be the ‘real’ transcription, which can be used either to modify the textual version, or to be rendered as such. In this phase, then, the transcriptionist has to use a specifically focussed and ‘constrained’ attention to a range of details, as specified in a version of the Jeffersonian conventions, treated as an analytic shopping list.

Actual transcription can be seen as a compromise between the two, balancing realist rendering and analytic utility, while still hoping to preserve a certain readability.

Both Ashmore & Read, and I myself, have used Bruno Latour’s concept of an ‘immutable mobile’

---

<sup>12</sup>For an incisive analysis of the visual identification of bird species using field guides see Law & Lynch (1988).

to characterize the functions of tapes and transcripts, but of course transcripts are not ‘immutable’ in a strict sense. I already referred to the possibility to use different versions of a transcript for different purposes, while the two versions of the laughter sequence transcribed by Jefferson, that I quoted before, demonstrate the fact that a transcript can be ameliorated by adding more details. The transcript by myself, that I quoted as an illustration, is a temporarily ‘finished’ product of a long period of successive ameliorations<sup>13</sup>

A transcript, then, is no more than a practically useful rendering of a recording of an actual interactional event. What is left of the original is limited to what can be heard and/or seen on the tape. The process of transcription reduces most of the actually hearable sounds to recognizable words in the standardized written version of the language used on the tape, while also allowing to add to this reduced version a number of symbols that evoke those aspects of the hearable sounds that have in the CA tradition acquired a status of potential interactional relevance, and thereby analytic utility. Furthermore, a transcript may serve – when given with a playing of the audio or video record – to instruct an audience as to what is there to be heard on the tape. In fact, when working on the transcript, the researcher may become only gradually aware of what there is to be heard<sup>14</sup>. The relationship between this after-the-fact constitution of the sense of an event, and the lived order of that event, is a problematic one. There are no final solutions to sense-making.

## References

Ashmore, M., D. 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-texte/3-00/3-00ashmorereed-e.htm>

Bogen, D. (1992) “The organization of talk”, *Qualitative Sociology*, 15: 273-96

Bogen, D. (1999) *Order without rules: Critical theory and the logic of conversation*, New York: SUNY Press

Clayman, S.E., D.W. Maynard (1995) “Ethnomethodology and conversation analysis”, In: P. ten Have, G. Psathas, eds., *Situated order: Studies in the social organization of talk and embodied activities*, Washington, D.C.: University Press of America, 1-30

Garfinkel, H., D.L. Wieder (1992) “Two incommensurable, asymmetrically alternate technologies of social analysis”, In: G. Watson, R.M. Seiler, eds. *Text in context: studies in ethnomethodology*, Newbury Park, etc. Sage: 175-206

Goode, D. (1994) ‘Construction and use of data in social science research’. In his *A word without*

---

<sup>13</sup>I provide one of these earlier versions in an appendix to this paper.

<sup>14</sup>David Goode (1994:150-62) provides some telling illustrations of these sense-making, sense-changing and sense-instruction possibilities.

words: *the social construction of children born deaf and blind*. Philadelphia: Temple University Press

Have, P. ten (1990) "Methodological issues in conversation analysis", *Bulletin de Méthodologie Sociologique*, Nr. 27 (June): 23-51 [also: <http://www.pscw.uva.nl/emca/mica.htm>]

Have, P. ten (1999) *Doing conversation analysis: a practical guide*, London, etc.: Sage Publications

Have, Paul ten (2001) 'Lay Diagnosis in Interaction' *Text* 21: 251–60

Heritage, J., J.M. Atkinson (1984) "Introduction", In: J.M. Atkinson, J. Heritage, eds. *Structures of Social Action: Studies in Conversation Analysis*, Cambridge: Cambridge University Press, 1-15

Hester, S., P. Eglin, eds.(1997) *Culture in action: studies in membership categorization analysis*, Washington, D.C.: University Press of America

Jefferson, G. (1985) "An exercise in the transcription and analysis of laughter". In: T.A. van Dijk, ed. *Handbook of discourse analysis*. London: Academic Press, Vol. 3, 25-34

Jefferson, G. (1990) 'List-construction as a task and a resource'. In: G. Psathas, eds. *Interaction Competence*. Washington, D.C.: University Press of America: 63-92

Latour, B. (1987) *Science in action: how to follow scientists and engineers through society*, Cambridge, Mass.: Harvard University Press

Law, J., M. Lynch (1988) "Lists, field guides, and the descriptive organization of seeing: Birdwatching as an exemplary observational activity", *Human Studies*, 11,271-304

Livingston, E. (1987) *Making sense of ethnomethodology*, London: Routledge & Kegan Paul

Lynch, M. (1993) *Scientific practice and ordinary action: ethnomethodology and social studies of science*, New York: Cambridge University Press

Lynch, M., D. Bogen (1994) "Harvey Sacks' primitive natural science", *Theory, Culture & Society*, 11,65-104

Lynch, M., D. Bogen (1996) "Methodological appendix: Postanalytic ethnomethodology". In their: *The spectacle of history: speech, text, and memory at the Iran-Contra hearings*, Durham & London: Duke University Press, 262-87

Psathas, G, T. Anderson (1990) "The 'practices' of transcription in conversation analysis", *Semiotica*, 78, 75-99

Sacks, H. (1984 a) "Notes on methodology". In: Atkinson, J.M., J. Heritage, eds. *Structures of Social Action: Studies in Conversation Analysis*, Cambridge: Cambridge University Press, 2-27



Sacks, H. (1984 b) "On doing "being ordinary"". In. Atkinson, J.M., J. Heritage, eds. *Structures of Social Action: Studies in Conversation Analysis*, Cambridge: Cambridge University Press, 413-29

Sacks, H. (1992) *Lectures on conversation*, 2 vols. Edited by Gail Jefferson with introductions by Emanuel A. Schegloff. Oxford: Basil Blackwell

## Appendix

Here is an earlier version of the transcript used in Illustration section

- A 'hh nou we zullen es kijken d't kan eh (.) eenvoudig (.) 'te zijn dat ze wat tekort aan bloed heeft ze is negen jaar 'hhh de leeftijden één jaar vier jaar en negen
- P (ja heb ik ook al)
- A ja tien elf zo'n beetje rond-tie tijd als ze een beetje uit gaan schieten 'hhh-dat zijn tijden waarop kinderen vaak een beetje bloedarmoede hebben
- P (jjjh) twee jaar geleden heeft ze 't ook gehad toen ook in september toen waren we bij de schoolarts en toen had ze ook bloedarmoede
- A hmm (..) 'k wee' niet of het wat is maar we kunnen even (prikken)