

# Open Research Online

---

The Open University's repository of research publications and other research outputs

## Reproducing or constructing? Some questions about transcription in social research

### Journal Item

How to cite:

Hammersley, Martyn (2010). Reproducing or constructing? Some questions about transcription in social research. *Qualitative Research*, 10(5) pp. 553–569.

For guidance on citations see [FAQs](#).

© 2010 The Authors

Version: Accepted Manuscript

Link(s) to article on publisher's website:  
<http://dx.doi.org/doi:10.1177/1468794110375230>

---

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's [data policy](#) on reuse of materials please consult the policies page.

---

[oro.open.ac.uk](http://oro.open.ac.uk)

**REPRODUCING OR CONSTRUCTING?  
SOME QUESTIONS ABOUT TRANSCRIPTION IN SOCIAL RESEARCH**

Martyn Hammersley  
The Open University

## **REPRODUCING OR CONSTRUCTING? SOME QUESTIONS ABOUT TRANSCRIPTION IN SOCIAL RESEARCH<sup>1</sup>**

Audio- and video-recordings are a major source of data in qualitative research today. There is now a substantial literature about the task of transcribing these recordings, though this mainly relates to linguistic and discourse analysis. In general, this takes the view that transcripts construct the talk or action that they portray rather than reproducing what is given. In this article I argue that while this is true in important respects, in that many decisions are involved in producing transcripts, there is also an important sense in which both the strict transcription of words used and the descriptions of speakers' behaviour are aimed at capturing something that exists independently of the transcription process. 'Construction' and 'givenness' are both metaphors and we need to be careful not to be misled by either of them.

Key words: data, evidence, transcription, constructionism, foundationalism

In recent years, there has been much discussion about the evidence that research can supply, and its capacity to make policymaking and occupational practices more 'evidence-based' (Trinder 2000; Thomas and Pring 2004). As part of this, concern has been expressed about the quality of the evidence used by social researchers in some fields, with particular doubts being raised about the rigour of qualitative methods. This has stemmed, in part, from the fact that, in the early stages of the evidence-based practice movement, the question of whether research itself is 'evidence-based' was addressed in terms of a gold standard exemplified by the randomised controlled trial. However, whatever the phrase 'evidence-based research' is intended to convey (Hammersley 2007), and however unconvincing the notion of a gold standard is, the nature of the evidence used by researchers, and how they produce and deploy it, is clearly an important methodological issue. Furthermore, it is one that has not been given the attention it deserves.

---

<sup>1</sup> An earlier version of this paper was given at the British Educational Research Association annual conference, Heriot-Watt University, Edinburgh, September, 2008. I am grateful to participants for their comments.

The focus of this paper is on a particular type of evidence, and one that is mainly employed by qualitative researchers: transcripts of electronic recordings. Use of such data is often regarded as more rigorous than reliance upon fieldnotes, in the sense that it provides a fuller and more accurate representation of 'what happened'. There are, however, some important questions about the nature of transcripts as data; and, in part, these centre on the meaning of the term 'data' itself. This concept is routinely taken for granted, but it involves hidden complexities.

### **What are data?**

From the point of view of what we might call a naïve foundationalist epistemology, the meaning of the term 'data' is relatively straightforward: crudely speaking, data are produced through objects in the world imprinting their characters upon our senses. This provides a foundation from which the scientist can then draw logical inferences to reach knowledge. Versions of this type of foundationalism, in more sophisticated form, can be found throughout the history of philosophy from Aristotle up to the logical positivism of the early twentieth century and beyond, albeit with variation in view about how the data are acquired, their character, and the form of inference that should be applied to them (see Losee 1993). Furthermore, foundationalist ideas have shaped much methodological thinking by social researchers, even including that concerned with qualitative work.

Within foundationalism, there are two analytically distinct senses given to the term 'data', and echoes of these continue to provide its main meanings. First, it refers to what is *given* to us from outside: that which is unmediated by us, or independent of us, and (so foundationalism assumes) can therefore be assumed to be beyond all possible doubt in terms of its validity. From this point of view, whatever is given simply *is* what it is; or, more precisely, it is what it *appears* to be.<sup>2</sup> The second meaning of the term is that data are the premises from which we draw logical inferences to conclusions; in other words, it concerns what we *take* to be given in the

inferences we make. This second sense of the term is functional, rather than substantive.

In the context of naïve foundationalism these two meanings fit together with no tension: we start with what is empirically given, and from this we then infer conclusions. Moreover, to the extent that the given is absolutely valid, and that we employ deductive inference (or some equivalent logical or mathematical calculus) to draw our conclusions, then those conclusions will be absolutely valid too. However, these two senses of the term ‘data’ have become separated as a result of criticisms of foundationalism.

One of the main arguments against foundationalism was that data are always theory-laden, or at least assumption-laden, so that they *cannot be* immediately given and automatically valid. Another criticism focused on the process of inference, pointing out that if this is ampliative – in other words, if it tells us something about the world beyond what is semantically implicit in the data itself – then the conclusion reached is always open to error. Moreover, this is true not just because a person drawing an inference can make a mistake in following the rules, which is also true of deductive inference and arithmetical calculation, but because the type of inference involved cannot *guarantee* the validity of the conclusion, in the way that propositional logic does.

However, criticisms of foundationalism have tended to be much more damaging for the notion of the givenness of data than for its role in the process of inference. It has come to be widely assumed, implicitly if not explicitly, that data do not need to be empirically given in order to serve their inferential function. And, indeed, without this relaxation in foundationalist dogma about what can count as data, the process of inquiry could not even get started. So, generally speaking, the notion of data as grounds for inference has been retained, while the idea of data as given has been abandoned. This is exemplified, for example, in the idea that data, and also ‘findings’, are constructed.

---

2 A key assumption of this sort of empiricism is that any mediation introduces error, or at least

The issue of transcription needs to be examined against the background of this decomposition in the meaning of 'data'. After all, it might be thought that, of all the kinds of data employed by social scientists, audio-recordings, and perhaps even transcripts of them, provide data that are in an important sense 'given'. Yet this idea has come to be widely rejected.

### **Transcription as construction**

Since the 1960s there has been an increasing tendency for qualitative researchers to use audio- and/or video-recordings, or transcriptions of these, as data – rather than fieldnotes. Indeed, reliance upon electronic recording and transcription has come to be taken-for-granted over the past few decades, so that fieldnotes are now treated in many quarters very much as a second-class kind of data, if their use is not ruled out completely. Indeed, some researchers have seen reliance upon electronic recordings and transcription as finally enabling human social interaction to be studied scientifically, since 'the data' are preserved and can be reproduced. This means that they are open to repeated analysis, and furthermore can be made available to readers of research reports so that analyses can be checked (and, in effect, replicated) by others. This is a view that can be found among conversation analysts (see Peräkylä 2003 and ten Have 2001:2). Of course, many qualitative researchers use electronic recordings and transcripts without viewing them in this way, treating them instead simply as a convenient alternative to fieldnote writing. From this point of view recording and transcribing, in contrast with fieldnote writing, are believed to provide enhanced detail and accuracy, without it being assumed that they supply a full or objective record of 'what happened', or that they are always the best source of data.<sup>3</sup>

There is now a substantial literature on issues surrounding transcription of audio- and video-recordings, though this literature does not seem to have been given much attention by most qualitative researchers. A great deal of it has been produced

---

the potential for it.

<sup>3</sup> It is not always feasible even to *audio*-record the processes of interaction in which one is interested, and the requirement to do so may exert undesirable constraints on what can be studied, where data collection can be carried out, and perhaps also on the length of the data collection period, given the sheer amount of data likely to be generated through electronic recording.

by those engaged in sociolinguistic research, including conversation analysis, though there are also discussions that relate to qualitative research more generally. A dominant theme in this literature is that transcription is a process of *construction* rather than simply a matter of writing down what was said.<sup>4</sup> What is meant by the constructional character of transcription is that a whole variety of decisions are involved, and that none of these is open to a single rational solution, so that there cannot be one correct transcription of any stretch of audio- or video-recording. For this reason, neither transcripts nor electronic recordings should be treated as data that are simply given, in an unmediated fashion, in the way assumed by foundationalism.

### **Decisions, decisions, decisions**

It is certainly true that a variety of decisions have to be made in the course of producing transcripts. This is highlighted by the fact that very different kinds of transcript are used by people working within different research traditions. One sharp contrast here is between the very detailed transcripts used by some sociolinguists and conversation analysts,<sup>5</sup> and the much less detailed ones employed by other sorts of qualitative researcher. There has even been a call for 'poetic transcription', which one advocate describes as 'the creation of poemlike compositions from the words of interviewees' (Glesne 1997:202).

The decisions involved in transcription include the following:

1. *Whether* to transcribe any particular audio- or video-recording, and if so *how much of it* to transcribe. Even conversation analysts do not always transcribe all of the recordings they make (Peräkylä 1993:206). Other qualitative researchers may be even more selective.<sup>6</sup>
2. How to represent the recorded talk. Here, there is variation according

---

4 See, for example, the early and influential article by Mishler (1991).

5 See, for example, the influential transcription system developed by Gail Jefferson (2004).

6 Here, I am thinking about 'how much' in diachronic terms: relating to issues like whether to start at the beginning of the recording, or instead when the 'real action' begins; whether to transcribe all of what is said or only those stretches of a recording that seem to be relevant to the research focus; and so on. I will address what we might call the synchronic interpretation of 'how much', for example whether to include non-verbal as well as verbal elements, under other headings.

to whether the emphasis is on capturing the actual sounds made or on identifying the words used and presenting these via traditional orthography. More specifically, there are decisions about whether to try to represent such features as intonation, pitch, amplitude, and pace of talk. Associated with this is the issue of whether to aim at capturing distinctive forms of language use, such as dialects. And, if pronunciation is to be represented, there is then the question of how to do this: whether through deviant spelling within traditional orthography or via a phonetic transcription system.

3. In multi-party talk, should there be an indication of to whom the speaker is primarily addressing the talk, where this is not indicated explicitly in what is said? While people sometimes address their speech to the whole company present, they do not *always* do this; sometimes they will pick out a sub-set of the group, even though other people will overhear what they say. And, occasionally, comments are exchanged between two or more members of a group in such a way as to avoid these being overheard by others, and this fact may be significant in understanding what is said. Some researchers using video-recording have sought to document to whom speech is directed by monitoring direction of gaze.
4. Whether to include non-word elements: such as back-channel noises (for instance 'uhuh'), laughs and other sounds that may be expressive. Also at issue is whether to report in-breaths and out-breaths, coughs, and other noises that may be regarded as non-expressive but still significant in some way. What noises are and are not intended to be communicative is not always clear; and, as Goffman (1959) pointed out long ago, information 'given off' can be as important as what is intentionally communicated.
5. Should silences and pauses be included in the transcript, should they be timed, and if so how? There are difficulties here about what counts as a silence or a pause, and about whether in timing them what is important is their actual length in seconds or whether they are likely to be perceived by speakers as long or short, significant or insignificant.

In one sense we might treat silence as simply the absence of talk, but there is a difference between this and notable silences or significant pauses. We also need to ask: significant or notable for whom?

6. Should we try to include relevant gestures and fine or gross physical movement, including for example what Peräkylä refers to amusingly as ‘ambulatory events’ (Peräkylä 1993:204), in other words walking about? Detailed information about these may be available via fieldnotes or where the events have been video-recorded, though not usually from audio-recordings alone.
7. There is also the issue of how to lay out the talk on the page in transcripts. While it is most common to use a playscript format of some kind, there are alternatives to this (see Ochs 1979). Moreover, even within the playscript format there are further decisions to be made. One concerns whether to treat all of what one person says, before another speaks, as a single continuous utterance; or to split and place on separate lines what might be interpreted as consecutive, distinct utterances by the same speaker, especially if these are separated by lengthy or significant pauses or silences. There is also the question of whether, and how, to indicate overlapping talk. It is important to note that overlapping talk is a distinct category from, albeit one that overlaps with, interruption.
8. There are alternative options in how to label the speakers. One possibility, rarely used, would be simply to number each utterance without indicating when the same person is speaking. In some ways, this might be in line with a postmodernist conception of identity as multiple and occasioned. More usually, labels are employed to indicate different speakers so that we can identify which utterances came from ‘the same’ source. However, it is important to remember here that speakership is not a straightforward matter, since people may speak on behalf of others or to a script written by others (Goffman 1981), and because all talk continually draws on other voices (Bakhtin 1981). Furthermore, any labels we give to speakers, beyond numbers or letters, may convey information about them, and this raises questions

about what information should and should not be included. For example, names can be taken to indicate gender (and inferences here can be false as well as true), and giving gendered names could imply that gender is the most significant factor in the interaction, which it may or may not be. Similarly, using role labels in transcripts, such as Teacher/Student, gives information and implies that all of a person's utterances were 'in role', and that these roles are the most important consideration in understanding what was going on.

9. When it comes to providing extracts from transcripts in research reports, there are further questions: where to begin and end the extract (for example, in the case of interview data, should the interviewer's questions be included?); whether to use the same transcript conventions as were used in the process of analysis, or ones tailored to the particular point being made in the text and/or to the audience; how many extracts to use in support of any point; what background information to provide about the speakers and the situation, and so on.

These decisions indicate the role of at least two sorts of constructive activity on the part of researchers and transcribers. First, there is *selectivity*: what to include and what to leave out, both in the full transcripts and in the extracts used in research reports. Secondly, there is the unavoidable use of cultural knowledge and skills by the transcriber to interpret and represent what is going on, for example in judging what is a significant pause. Indeed, it is very important to emphasise that what is involved in transcription is not a matter of writing down sounds in some etic fashion. At the most basic level, the sound stream must be 'broken up' into identifiable heard words belonging to a particular language. And because of variation in the pitch of voices, in pronunciation, etc, there will not be any simple correspondence between the sound in etic terms and the words recognised. More than this, what we hear as transcribers are utterances of particular kinds, exemplifying particular actions. It is *within* this kind of understanding of what is being said and done that we identify particular words. Nor is it a simple matter of turning heard words into written words on a page: generally speaking the aim is to do this *so as to convey what was being said, how, with what emphasis and import, and so on*. And this requires more than just knowledge of the

language, narrowly understood in terms of a sound system, lexicon, and grammar. So, the identification and representation of the words being spoken depends to some extent upon our ability to understand what the person might be meaning by what he or she says. We will recognise words much more easily when they seem to be part of a meaningful statement than when they are part of an apparently meaningless one. To underline the point being made here, we do not first transcribe the sounds and only then try to identify the meanings, we ‘hear’ sounds and meanings simultaneously in the course of transcription. It is also true that *what* we transcribe, and to some extent *how* we transcribe it, reflects substantive assumptions (about human beings and their social institutions) and methodological ones too (about how best to describe and explain social phenomena).

### **Dangerous delusions**

Given the role of all these decisions, and especially the reliance upon interpretation, it is very important to recognise the constructive work that goes into transcription. But I also want to suggest that there are serious dangers with the predominant emphasis on the constructed character of transcripts that is now all-pervasive in the literature on this topic. There is a slippery slope from recognition that decisions and interpretations are necessarily involved in transcription to the conclusion that the data are created or constituted by the transcriber rather than representing more or less adequately 'what occurred'. In effect, this leads to a radical epistemological scepticism that is self-undermining. There is a tendency in this direction in some influential methodological writing about qualitative research generally today. In the case of transcription, this epistemological radicalism takes the constructed character of transcripts to indicate that they are fictions, that they are *made up*. For example, in an article on the 'experiential text', Norman Denzin writes:

The “original” voices of individuals in a field setting, and the intentions behind those voices, can never be recovered. There are no original voices, for every instance of a voice being heard is an original, a new hearing, a new voice speaking (and hearing) its mind. Every transcription is a re-telling, a

new telling of a previously heard, now newly heard, voice (p14).

Earlier in the same article Denzin claims that ‘the worlds we study are *created* through the texts that we write’ (p9). In my view, this kind of constructionist scepticism is misconceived and misleading (Hammersley 2008).

Central to much discussion of the constructed character of transcription is a rejection of any notion of givenness in relation to data. However, we need to remember that we are dealing with metaphors when we talk about construction versus givenness, and that neither of these metaphors on its own captures perfectly what is involved in transcription. This means that we should use them with some circumspection. I suggest that *both* concepts capture some aspects of what is involved, and by refining our understanding of what they do and do not represent adequately we can get a better grip on the phenomenon and refine these concepts. Much of our thinking is metaphorical, in the sense that we always try to understand the things we find puzzling in terms that we have used successfully in understanding other things. And in that process we hope to develop concepts that will be more adequate to the new task, reaching a point where they *do* capture effectively most or all of the relevant aspects of that phenomenon, for current purposes. So our need to rely upon metaphors does not itself mean that our thought is essentially fictional, only that we use what we assume we already know to try to get a grip on what we do not yet understand. Moreover, all knowledge is knowledge of answers to particular questions, it does not offer a totalising grasp on the very nature of the world. So, what I want to suggest here is that, while the notion of givenness as it is used in naïve foundationalism is misleading, this metaphor does nonetheless still tell us something important about transcripts as data; and something that is lost if we take the notion of construction too far.<sup>7</sup>

At this point, I think we need to explore in more detail exactly what the transcription of audio-recordings involves. What we routinely refer to as transcripts are not homogeneous in character. One element is what I will term ‘*transcription in*

---

<sup>7</sup> If there is a need to appeal to influential theorists in support of the notion of givenness we could pick Gadamer and Heidegger: neither was an extreme constructionist or sceptic, both believed

*the strict sense* or *'strict transcription'* for short.

### **Strict transcription**

This relates primarily to those parts of a transcript where the words heard on a recording are written down by the transcriber in the script of some language.<sup>8</sup> Besides the transcription of words via standard orthography, as noted earlier there are also sometimes attempts to represent pronunciation, either through deviant spellings of words or via use of the international phonetic alphabet, or some other system. This too is part of strict transcription; though I suggest that it is secondary to the task of writing down the words spoken. It is derivative in the sense that it is the words spoken that are the starting point, only subsequently do we try to represent their pronunciation and other features of *how* they were spoken. Much the same is true of attempts to represent other features of speech.

Even strict transcription is constructional, in the sense that it relies upon the transcriber's linguistic abilities. Moreover, as I indicated earlier, transcription cannot be viewed entirely in linguistic terms, narrowly defined – it necessarily also depends upon the transcriber's ability to use her or his knowledge of the language and culture to make sense of what people are, or could be, saying and doing. There may even be a contribution from general knowledge about human needs, practices, and tendencies. In short, what is involved is what has long been referred to as *Verstehen* (see Martin 1999).

At the same time, we should recognise that the kind of listening that is required in transcription is different from that which we do during the course of our

---

that understanding involves contact with something beyond our own sense-making capabilities, though both sometimes framed this in rather mystical terms. On Gadamer, see Wachterhauser 2002.

<sup>8</sup> Of course, not all languages have a written form. Moreover, some of the literature on transcription has been concerned with the extent to which the written form of a language, which is usually based on the dominant spoken variety, is capable of reproducing subordinate varieties accurately. What is crucial in the case of strict transcription, above all, is whether the words used can be spelt in the language, and thereby made available in written form. It is also worth remembering that transcription is sometimes the first stage in a process of translation between languages, where the language used by the people being studied is different from the language in which the analysis is to be carried out.

ordinary interactions. In producing strict transcriptions we are using whatever understanding we gain of the stretch of social interaction concerned, of what people are doing and why, *in order to identify and write down the words used, to recognise significant pauses, etc.* By contrast, in ordinary interaction, most of the time, we are not concerned with the particular words used, or with pauses, for their own sake, but rather with understanding what is being said, what its implications are, what response needs to be made, and so on.<sup>9</sup>

Moreover, while strict transcription is constructional in the respects just indicated, it is also necessary to recognise that it involves an attempt to capture features that are in an important sense given. This argument can be taken back to the very process of recording. Audio-recordings preserve some of the sounds from the occasion that was recorded, and in particular (usually) those sounds within the frequency range relevant to most human speech. Of course, audio-recording machines are human products, and so it is no accident that they record this frequency range. Furthermore, in order to record what will be recognised as talk they must be placed in a time and place where talk occurs and can be picked up by the microphone. However, once placed and switched on, *what particular sounds* these devices preserve on record is not determined by the researcher: the recorded sounds are a product of what happened in the situation at the time of the recording (within the limits of what is recordable by the device concerned).<sup>10</sup> So, while there is activity on the part of the researcher in setting up the recording process, there is an important sense in which the sounds recorded are *given* rather than created by her or him.

The same point applies to the process of transcription itself. When we listen to an audio-recording, what we hear is given, in the sense that we are seeking, first and foremost, to write down the words heard. And *hearing* particular words and utterances is not the same as *imagining them* or *inventing* them. We can *mis-hear*

---

9 What is demanded by social research, here and elsewhere, is a slowing down of, and the giving of reflexive attention to, normal processes of perception and cognition; in order both to increase accuracy and to provide for the communication of evidence.

10 If the event is a research interview, then the researcher will have played a role in determining what is said. Similarly, there may be reactive effects in observation. However, this issue of reactivity is a separate one, it is not intrinsic to audio-recording and transcription; there can, after all, be secret recording of what happens in situations where a researcher is not present, and where the fact that recording is taking place is not known.

particular words, which indicates that what is hearable on a recording is not an infallible representation of itself, yet this also highlights the fact that there is a genuine sense in which what is heard is *there* rather than *made up*. To labour the point, there is an important difference between a transcriber writing down what people said (in a context where this can be heard relatively straightforwardly) and a novelist inventing what her or his characters are going to say. To over-emphasise the similarities here, under the influence of constructionism, is to neglect the even more important differences; along with the need to find a way of thinking about givenness that escapes the problems associated with how it is treated by foundationalism.

We should also note that there is a sense in which the metaphor of construction itself relies upon some notion of the given. All construction work employs resources and tools that are taken for granted. So, both the transcriber seeking to write down the words used, *and* the novelist seeking to invent dialogue, take various things as given. However, they treat *different* things in this way because they are engaged in different activities. The transcriber treats the words heard as given, even though he or she may revise judgement about what was heard in any particular instance. The novelist does not have a straightforward equivalent to this (unless genuinely hearing the voice of a muse!), but, like the transcriber, he or she takes the grammar of the language as largely given, while at the same time relying much more on ideas about what a particular character, or any character, *would* say or *might* say in the circumstances. Furthermore, it is worth pointing out that inventing dialogue is parasitic in developmental terms: we could not do this if we had not already heard people speaking and were not already able to write down what they had said.

So what I have referred to as strict transcription involves constructional work but also takes certain resources as given, notably the words that are audible in a recording. It is worth noting that if we were to raise doubts about the existence or possibility of hearing the words other people are speaking, in the manner that Denzin seems to do, we would have little or no grounds for beliefs about anything, so basic is this assumption to our whole existence. A thorough scepticism would result; and this undermines itself, since it leaves us no grounds on which to stand in order even to raise doubts in the first place. Doubting anything always involves taking something

else for granted, and there are few things that are more taken for granted by most human beings than that they can hear others speaking and know what they are saying.

What I have referred to as strict transcription does not exhaust what is routinely included in the documents we refer to as transcripts. Also involved are descriptions, and I now turn to a discussion of these.<sup>11</sup>

## **Description**

Transcripts usually involve at least some descriptions: of who is talking, how they are speaking (seriously, ironically, jokingly, sarcastically, etc), non-verbal behaviour, contextual information, and so on. This is true when transcribing audio-recordings, but description usually takes on an even more important role in transcriptions of video-recordings. While there are systems designed to provide for strict transcription of non-verbal behaviour, these are by no means always used (see Birdwhistell 1973; Goodwin 1981; Kendon 1990).

In the case of transcribing audio-recordings, the issue of whether non-verbal vocal expression, like laughter and crying, should be dealt with via description or represented by means of some form of phonetic transcription has been an issue in the literature. Jefferson (1985) and Hepburn (2004) have insisted on the importance of phonetic transcription. Ashmore et al (2004) point out, against this, not only that there are considerable descriptive resources available within the English language distinguishing kinds of laughter and crying, but also that there is no direct or immediate relationship between the sounds made and what sort of laughter or crying is taking place, or even whether this is what is actually going on.<sup>12</sup>

Descriptions are very different in character from strict transcriptions: they do

---

11 It is perhaps worth pointing out that even fieldnotes may well include an element of strict transcription, where people's words are reported in what is intended to be a verbatim manner. But, of course, the balance between this and description is usually the reverse of that found in transcripts: the stretches of strict transcription are likely to be shorter, as well as being less accurate.

12 It should be said, though, that most descriptions of laughter in transcripts do not go much beyond such brief, and mundane, labels as 'laughs'.

not involve identifying words spoken but, rather, characterising actions in terms of categories. Even more obviously than with strict transcription, they involve processes of interpretation and judgement. This derives from the fact that there are many different true ways of describing the same phenomenon, because in effect a description always amounts to answers to some explicit or implicit set of questions about the phenomenon concerned. By contrast, there is only one true way of transcribing a stretch of talk in the strict sense of writing down the words – leaving aside the issue of alternative ways of indicating pronunciation, which I have already suggested is derivative. So, in descriptions we *use* language to identify objects and characterise their attributes, rather than re-presenting language features used by others; and how we select the objects to be described, how we formulate their character, what out of all the various attributes we *could* ascribe to them we actually do, and so on, depends upon the framework of relevance defined by the questions our description is seeking to address. As a result, a large number of descriptions can be provided of any phenomenon, each responding to different questions. While some of these will overlap in what they cover, to varying degrees, others may involve little or no overlap at all.<sup>13</sup>

So, description is always in terms of a particular framework that picks out some features of the objects being described as relevant and ignores others; and, of course, what is heard and seen must be interpreted in terms of that framework. Given this, should we conclude that description is entirely a matter of construction? My answer to this is no. There is an important element of givenness even in descriptions. Within the field defined by the framework adopted, we perceive X to be happening rather than Y or Z, so that what we subsequently describe *will have been* constrained by evidence about what is going on in the situation. And this will be true even if we *misperceive* the scene, since this will be a misperception in some particular respect, not a hallucination or a dream.<sup>14</sup> Here again, there is a difference between setting out

---

13 It might be argued that actions are a peculiar type of object, one for which there is only one true characterisation in each case, as determined by the actor's intention. However, the role of intentions in characterising actions is controversial, and we should note that even if we take intention to determine the type of action involved, intentions usually form part of nested structures: we do A in order to achieve B, in order to gain C, and so on. Given this, there will still be a range of alternative descriptions of any action: as A, B, or C, etc.

14 I will not address the question of whether all perceptions could be hallucinations or dreams because if they are I would have to assume that I am only hallucinating or dreaming that I am writing

to write an account of what happened and making up an imaginary story, one which the concept of mis-perception highlights. And it is a difference that we have good reason to preserve, not only because it underpins the whole research enterprise but also because it is essential to us in our everyday lives. While description involves a different kind of construction process from strict transcription, and may involve more serious threats to validity, it too relies on what we can take as given in our experience of the world. And while we cannot assume that our perception or judgement of what is given is infallible, or that what we perceive is always what exists, neither should we assume that what we see or hear must always be false or that it necessarily amounts to creation out of nothing. We should not do this because there are no good reasons for doing so.

What is involved here is a revision of the notion of givenness built into naïve foundationalism.<sup>15</sup> It is not a matter of the given presenting itself to our senses in a manner that immediately indicates its character with certainty. Indeed, the character of the given is not what is at issue but rather what it can tell us about relevant features of the phenomena in which we are interested. In this context, we should note that even when, as researchers rather than transcribers, we read those parts of a transcript that are strict transcription, we necessarily generate from these a conception, or description, of what is going on: not just what is being said but what is being done, what it means, why it is being done, what implications it carries, and so on. In other words, in order even for the strict transcription components of transcripts to function as data – in the sense of as grounds for inference – they must be turned into descriptions. Another way of putting this is to say that using data involves inference from signs.<sup>16</sup> The metaphor of construction highlights the role that inference plays here. But the concept of sign also carries with it the implication that two things are given: the material that is the vehicle for the sign *and* that to which the sign points.

---

this paper, and so there would be no need for me to worry about what doubts my audience might have about my argument!

15 There may be resources available for developing this further in Peirce's discussions of firstness, secondness and thirdness, along with his theory of signs; and also in Husserlian phenomenology.

16 This is a quite ancient way of understanding the notion of evidence that goes back at least to Aristotle (see Allen 2001).

## Conclusion

What we have in transcripts treated as data, then, are not representations of sounds recorded but rather of words heard that are components of utterances that are taken to represent actions of particular kinds; along with descriptions of relevant associated features that are designed to aid the process of inference. Both transcripts themselves and what they are taken to show, are indeed constructed, in an important sense, but they also rely upon what is given when we listen to or watch recordings. And the phenomena we draw inferences about are also treated as given independently of us: signs are taken to point to the existence of that which they signify.

So what conclusions can be drawn from my discussion? These fall into what we might broadly call theoretical and practical categories. In this article I have concentrated on theoretical issues raised by the now commonly accepted view that transcripts are constructions. I have argued that this can be misleading. It can result in an explicit or implicit rejection of givenness that is mistaken. I suggested that we must remember that the notions of construction and givenness are metaphors, and we should consider in what senses each captures the process of transcription and what it misses. Both metaphors have something important to tell us, but both can lead us astray.

I also pointed out that the transcripts we produce and use as part of research are not homogeneous: they are composed of both strict transcription and description. However, for transcripts to serve as data for analytic purposes, they must all, even the strict transcription components, be used as a basis for generating at least implicit descriptions. In other words, they must be interpreted as representing some set of actions and/or events, some sort of scene involving people with intentions and motives, who are participating in practices and institutions of various kinds, and so on. *It is such descriptions that are the data of social science, from which inferences to conclusions and findings are made.* This point counts against what Malcolm Ashmore has referred to as tape or transcript fetishism, which he ascribes to some conversation analysts; and also against the idea that all social life must be understood as discourse. However, it should not lead us to the kind of extreme constructionism or fictionalism

apparently adopted by Denzin and others, where research is equated with literature or art (Hammersley 2008). After all, description involves using what seems to be given in order to characterise what is there, from the point of view of some set of questions. In short, it employs inference from signs to conclusions about what is going on and why.

What is involved in transcription, and in the whole activity of research, is a slowing down and reflexive re-routing of a process that operates much more rapidly in ordinary social interaction, where we 'hear' what people say; and 'hearing' means understanding what they mean, this necessarily depending upon some grasp of why they are saying and doing what they are, what sort of response may be required, and so on. Electronic recording preserves a record of some of the data on which as participants or observers we might make sense of what is happening. Transcribing it involves presenting the words we can hear in written form, and providing descriptive resources for interpreting them in a much more deliberate fashion for the purposes of social science. However, even strict transcription cannot be carried out on its own: it relies on *Verstehen*, even though it is designed to allow a more reflective process of interpretation that will facilitate the building of social scientific knowledge. So, the recognition of words in what I called strict transcription is necessarily guided by interpretations of what is being communicated and what is being done.

As regards the practical implications of my argument, one point to be made is that great care is needed in transcription: it is easy for errors to creep in (Poland 1995) and these can lead to false inferences. In strict transcription we need to try to ensure that we are identifying the words, and/or phonetic characteristics, accurately. In terms of description, we need to be clear about the relevancies we are using in deciding what to describe and how to describe it. This will depend to some extent upon our research questions, though of course these need not be fixed or very closely defined, especially at the beginning of the research, so that they may leave open a wide range of potentially relevant matters. It is also important to try to ensure that, within the relevance framework adopted, we have included in our descriptions everything that might be relevant to understanding what is going on, from the perspective we have adopted, and describing it accurately.

Equally important, we must not treat transcripts as sacred and infallible texts. Even strict transcription of the words spoken does not guarantee to tell us what someone was meaning to say or what they were doing. We have to *interpret* the words, and in doing so we will and should draw on our experience of observing the events concerned (where available), fieldnote descriptions of them, general background knowledge, and so on. While we must be careful not to over-interpret what people say – in the sense of engaging in speculative ascription of intentions and motives, social functions, etc. – what words they literally pronounce does not in itself tell us what they were intending to communicate or do. An accurate strict transcript, whether based on repeated listening to audio-recordings or produced ‘live’, preserves some of the evidence in a relatively concrete form that may be necessary for us to assess, and re-assess, our inferences. It can do no more, and should do no less, than this.

Another practical conclusion is that different sorts of transcript, as regards both strict transcription and description, may be required depending upon the purpose for which the transcript is to be used. Most obviously, what should be included in a transcript, and how this should be represented, will vary according to the nature of the investigation. For example, it will be different if we are seeking to analyse turn-taking in conversations from if we are using data for the purposes of understanding social strategies employed by parents to control their children.

However, it is also important to recognise that which is the best form of transcription may well vary over the course of any single inquiry. Indeed, as we have noted, at the beginning, the researcher may be unclear about what would and would not be relevant to include in transcripts for the purpose of facilitating the analysis. And it is very important here not to allow some single transcription scheme to determine this. The focus must be on producing relevant and accurate descriptive material with which to try to answer one’s research questions. Transcripts are only tools to be used in this work, they never exhaust the potential data; nor, for that matter, do the audio- or video-recordings. Towards the end of a research project what is and is not relevant may be much clearer, and a rather different sort of transcript may

be most useful in facilitating the analysis. This parallels the situation in quantitative research, where it may be necessary to produce different sorts of table at different stages of inquiry.

It is also important to consider how extracts from the transcribed data should be presented in research reports, since here we must think about communication with others not just about reaching sound conclusions ourselves. Indeed, the task is precisely to convey what we take to be the sound conclusions we have reached, and to do so in a way that aims to be persuasive, within the legitimate requirements of this, while at the same time allowing others to make judgements about the validity of our claims. There are conflicting considerations here. On the one hand, we need to make the nature of the evidence we are using clear to our audience. And one consideration is that very complex and detailed transcripts may not serve us very well. For many audiences such transcriptions will be unreadable – this is true of the sorts of detailed transcripts produced both by conversation analysts and by some linguists. On the other hand, we ought to want to provide readers with data and evidence in a form that allows them to at least consider whether alternative interpretations from those put forward by the researcher would be plausible. And this may require that we put in more details than those that are directly relevant to the specific knowledge claims we are making. This means that readers may need to learn to interpret and tolerate more complex transcriptions than those they are familiar with. Of course, there is no possibility of including in a transcript everything that *could* possibly be relevant. While the researcher should identify and assess plausible alternative interpretations, and provide transcripts that allow these to be assessed, he or she will not be able to anticipate all alternatives. And, as already indicated, too much detail will obscure relevant data and how it relates to the knowledge claims being made. Some trade off between these alternatives – between very inclusive transcripts, on the one hand, and highly selective ones, on the other – is required.<sup>17</sup>

The aim of this paper has been to highlight some of the complexities involved in transcription, and the importance of thinking about these in the course of doing

---

<sup>17</sup> Of course, there is the possibility of making detailed transcripts, and even the recordings themselves, available via archives, though this is not entirely straightforward for ethical and other

qualitative research. This is part of the larger issue of what counts as data or evidence in research, and the requirements relating to this. My main point is that here, as elsewhere, we must take care not to become prisoners of particular metaphors, whether that of givenness or construction, but rather to assess what each can teach us about the process we are trying to understand.

## References

Allen, J. (2001) *Inference from Signs: Ancient debates about the nature of evidence*, Oxford, Oxford University Press.

Ashmore, M. and Reed, D. (2000) 'Innocence and nostalgia in conversation analysis: the dynamic relations of tape and transcript', *Forum: Qualitative Research*, 1, 3 (on-line journal). Available at:

<http://qualitative-research.net/fqs/fqs-eng.htm> [accessed 15.02.08]

Ashmore, M., MacMillan, K., and Brown, S. D. (2004) "'It's a Scream": Professional hearing and tape fetishism', *Journal of Pragmatics* 36(2), pp349-374.

Bakhtin, M. M. (1981) *The Dialogic Imagination: Four essays*, (ed. M. Holquist), Austin, University of Texas Press.

Birdwhistell, R. (1972) *Kinesics and Context*, Harmondsworth, Penguin.

Bucholtz, M. (2007) 'Variation in transcription', *Discourse Studies*, 9, 6, pp784-808.

Denzin, N. K. (1995) 'The experiential text and the limits of visual understanding', *Educational Theory*, 45, 1, pp7-18.

Glesne, C. (1997) 'That rare feeling: re-presenting research through poetic transcription', *Qualitative Inquiry*, 3, 2, pp202-21.

---

reasons. On the issues surrounding archiving and reworking qualitative data see Heaton 2004 and the

- Goffman, E. (1959) *The Presentation of Self in Everyday Life*, Harmondsworth, Penguin.
- Goffman, E. (1981) *Forms of Talk*, Oxford, Blackwell.
- Goodwin, C. (1981) *Conversational Organization: Interaction between speakers and hearers*, New York, Academic Press.
- Hammersley, M. (2007) 'Is it even true that research is evidence-based?', unpublished paper.
- Hammersley, M. (2008) *Questioning Qualitative Inquiry*, London, Sage.
- ten Have, P. (2002) 'Reflections on transcription', *Cahiers de Praxématique*, 39, pp21-43. Available at (accessed 27.08.08):  
<http://www.paultenhave.nl/Transcription-rv2.pdf>
- Heath, C. (1997) 'The analysis of activities in face to face interaction using video', in Silverman, D. (ed.) *Qualitative Research: Theory, method and practice*, London, Sage.
- Hepburn, A. (2004) 'Crying: Notes on Description, Transcription, and Interaction', *Research on Language & Social Interaction*, 37, pp251-91.
- Jefferson, G. (2004) 'Glossary of transcript symbols with an introduction', in Lerner, G. H. (ed.) *Conversation Analysis: Studies from the first generation*, Amsterdam, Benjamins.
- Kendon, A. (1990) *Conducting Interaction*, Cambridge, Cambridge University Press.
- Kitzinger, C. (1998) 'Inaccuracies in quoting from data transcripts: or inaccuracy in

quotations from data transcripts', *Discourse and Society*, 9, pp136-43.

Losee, J. (1993) *A Historical Introduction to the Philosophy of Science*, Third edition, Oxford, Oxford University Press.

Martin, M. (1999) *Verstehen: the use of understanding in the social sciences*, Chicago, Transaction Books.

Mishler, E. G. (1991) 'Representing discourse: the rhetoric of transcription', *Journal of Narrative and Life History*, 1, 4, pp225-80.

Ochs, E. (1979) 'Transcription as theory' in Ochs, E. and Schieffelin, B.B. (eds) *Developmental Pragmatics*, London, Academic Press.

Peräkylä, A. (1997/2003) Validity and reliability in research based on tapes and transcripts. In Silverman, D. (ed) *Qualitative Analysis: Issues of Theory and Method*. London: Sage, pp. 201-220. Revised edition 2003. Reprinted in Seale, C. (ed.) *Social Research Methods. A Reader*. London: Routledge. 2003

Poland, B. D. (1995) 'Transcription quality as an aspect of rigor in qualitative research', *Qualitative Inquiry*, 1, 3, pp290-310.

Poland, B. (2002) 'Transcription quality' in Gubrium, J.F. and Holstein, J.A. (eds) *Handbook of Interview Research*, Thousand Oaks, Sage. Reprinted in Holstein, J.A. And Gubrium, J.F. (eds) (2003) *Inside Interviewing: New Lenses, New Concerns*, Thousand Oaks, CA, Sage.

Thomas, G. and Pring, R. (eds.) (2004) *Evidence-based Practice in Education*, Maidenhead, Open University Press.

Trinder, L. with Reynolds, S. (eds.) (2000) *Evidence-Based Practice: a critical appraisal*, Oxford, Blackwell Science.

Wachterhauser, B. (2002) 'Getting it right', in Dostal, R. J. (ed.) *The Cambridge Companion to Gadamer*, Cambridge, Cambridge University Press.