

Moving Performance to Text: Can Performance be Transcribed?

Edward L. Schieffelin

Introduction

Transcription is the production of a written version of material originally presented in performative form. In its highest aspiration, it would be the attempt to record transparently and objectively in writing every significant detail of a performance, including the tone and emphasis, pacing and synchronization, and momentum and intensity of events, in the order in which they occur. A transcript differs considerably from a script—even a script with actor's and director's notes. A script is an outline, a prescriptive guide, for the production of a performance—for what a performance may be. It mandates an indefinite number of possible performances. A transcription, by contrast, is a record of a specific performance event. It is, in this sense, a kind of historical document whose purpose is to record every detail of something that has already actually occurred. One might say that a script prescribes the performance, the performance interprets these prescriptions in playing them out, and the transcription attempts to detail the result in writing. Because a transcription records actual rather than prescribed events, it aspires to be the ultimate form of entextualization of performance.

It is interesting, then, given the quality and uniqueness of famous performances, that transcription is regarded as an inferior genre among the many literatures that relate to performance. The bookstores of the National Theatre, the Barbican, or the Globe bulge with scripts, commentaries, and histories relating to great performance pieces, but no transcripts. My main purpose in this paper is to explore the production and use of this orphaned form of performance-related text and to consider what kind of a representation of a performance a transcript is.

In what follows I will restrict myself largely to discussion of the transcription of verbal and aural components of performance materials. This restriction is not to slight the special complexities of visual transcription but for simplicity, because encompassing the special issues of visual transcription would not add to the general points I wish to make.

What is special about the transcription of verbal art performance? What are we trying to get at in moving a performance to text in this way? To start with, transcription entails much more than the transfer of the verbal content of a performance into text. For me, and I speak as a cultural anthropologist rather than a historian or theorist of the arts, a performance is first and foremost a living event. When it is over, it is gone. There may be another similar performance tomorrow if it doesn't rain, but that is another performance.

The transcription of a performance, if it is to capture its performative qualities, is always more than the entextualization of the verbal or musical content. What seems to me to be of central interest in a performance, especially when studied cross-culturally, is its *mode of process*: how it works, how it brings about the effects that its participants intend. While this is not the only focus in performance studies, all other concerns—of history, genre, form and practice, aesthetic sensibility and theory, production complexities, dramatic means, and so on—all converge around, or ultimately refer to, the strategic processes by which a performance works. What does it mean to say, in a given culture, that a performance works? For scholars of non-western literatures, as well as for anthropologists, this issue begs the question about the culture within which the performance is embedded, from which it arises as a creative entity, and which it in turn addresses. For an anthropologist, the approach to all these questions, including those about performance, must be ethnographic.

An ethnographic perspective views any performance, in the first instance, as a social event. From this perspective, a performance works only because it has a relationship with (and an effect upon) others: in effect, an "audience." Thus, insofar as a performance is addressed to a responsive audience, it must be investigated as, in an important sense, co-created in its working with that audience. This immediately problematizes the boundaries of performance for transcription—for if performance is a responsive (rather than purely presentational) genre, must the audience be considered part of the performance? This may not seem much of a problem in the context of ordinary western art theater, where by convention the audience retains a relatively quiet, passive-attentive posture; it is easy enough to believe the performance is restricted to the stage. But such is not the case for many forms of cabaret performance or modern experimental theatre—let alone for

highly interactional genres such as preaching in Black American churches, or many highly performative ceremonial and ritual contexts in non-western settings. Is the “audience” here part of the performance or part of the contingent context? This is one of the questions the making of a transcript is forced to address.

The situation is further complicated for anthropologists when a genre that is considered a “performance” by an outside observer (a ritual, for example) is seen by its practitioners as quite the opposite. The rather dramatic Bosavi spirit séance, whose transcription I will be discussing shortly, was considered by its indigenous audience to be the very opposite of a “performance” (in the western theatrical sense). Either the spirits who came up and spoke were really present or they were not. In fact, those gathered about the medium were highly alert to the possibility that the medium might be “acting,” in the western sense. If so, if the medium and not the spirits produced the spirit voices, then the séance was fraudulent and the audience was being deceived. Given that séances were important means for dealing with such politically explosive issues as illness, death, and witchcraft, this kind of deception could be a very serious matter.

In such a situation, I deem it impossible to find a universal, properly bounded definition of performance that would indicate the domain we are to transcribe. My approach is not to engage with this problem on definitional grounds but to seek a working perspective. For the purposes of my own research, I would simply observe that (1) however one characterizes “performance” it is a contingent event, and (2) any human social event that involves expressive and communicative aspects can be usefully subjected to performative analysis—that is, analysis in terms of its performative aspects.

It follows that a transcription of a performance event may legitimately (and, I would assert, should in principle) incorporate audience activity and any other occurrence that may affect the expressive structure or content of the event. In the Bosavi spirit séances I studied, which took place at night, even small noises of the darkness outside the longhouse sometimes had dramatic consequences for the way a séance was carried out—and had to be incorporated in the transcript. It is a “performative event” then, that we are to transcribe—and that includes anything that affects or gives context to the performative activity, whether or not, by western standards, it might count as part of a “performance” (Schieffelin 1998).

Once we decide upon the performative event that we wish to transcribe, a new set of issues present themselves. Live performances themselves are hardly ever transcribed as they occur. The content and flow of an ongoing performance is much too rich and fast-moving to be written down simultaneously. It must be rendered in a form that can be slowed

down, repeatedly viewed, and minutely considered—in short, in a form that holds still long enough to be converted into text. The process of transcription thus begins with the making of a recording—on tape or video—of the performance event, and it is the recording, not the performance itself, that is transcribed.

Recording

Making a recording is rarely a fully straightforward process. Quite apart from the usual technical issues of making sure the recorded material is clear and audible enough to be transcribed, there is the problem of covering the performance space. The space of many indigenous performance events may not be well defined, may change in size, density, and orientation, and may travel from place to place. Indeed significant aspects of the performance may occur in several places at once with considerable overlap. How to record this for problem-free transcription is challenging and calls for careful planning and (ideally) prior experience with the kind of performance to be recorded.

Besides this, performances often take place in a number of registers among different subgroups or interactants and via a number of different media. For purposes of completeness, the researcher usually has to simultaneously track more than one of these, and recording equipment must be set up accordingly.

In addition to all of this, the presence of a recordist and his or her equipment and possible assistants must be factored in. There are some who argue that if the recordist remains sufficiently unobtrusive the performance will unfold almost exactly as it would in a “pristine” state—that is, in the absence of equipment and recordist. Although this may sometimes seem to be the case, it is best to avoid this assumption. In my experience, it is impossible for a recordist and his or her equipment to completely escape the awareness of the participants and not affect the performance in some way. It is best to make a virtue of one’s presence by acknowledging that the performance one is recording is one at which there is a recordist present. How the recordist is accommodated by the performance, then, becomes an interesting part of the way the performance takes place.

It should be clear by now that recording for transcription unavoidably will involve a number of compromises depending on the interests of the researcher and the contingencies of the particular situation. Without expatiating further on these problems, it is of interest to see how they were worked out in a particular case. I will describe my own experience with

recording and transcribing Bosavi spirit séances not just because it is familiar to me but because I believe that it represents an example of particular complexity that involves many of the kinds of choices a researcher of performance is likely to encounter.

When I began studying Bosavi spirit mediumship, I was impressed with the lively, highly interactive, and occasionally raucous nature of the séances. The fast-moving nature of these performances, and their detailed developmental structure, required recording and transcription if they were to be given detailed analysis.

Fortunately for me these performances took place in nearly complete darkness so that the visual component of the activity was insignificant and, I hoped, could safely be ignored. I also hoped my presence in the darkness would be forgotten and that the performance might truly proceed as if I were not there. But that was not to be. Every now and then some solicitous soul would ask if I was OK or still listening, or comment on my presence to others. The worst moments came from my indigenous research assistants who, in moments of intensity, occasionally called out to people not to all speak at once lest the conversation prove impossible to transcribe. These incidents were few and far between, and most of the time people seemed to ignore my presence—but never enough for me to feel that my presence was fully irrelevant to the performance.

If the darkness removed the necessity for dealing with the complexities of the visual aspects of performance, it created other problems. A Bosavi spirit séance takes place in a communal longhouse. The spirit medium lies down on a sleeping platform and other members of the community group themselves around him—some sitting close up, others just behind them, while others sit further in the background. Those closest to the medium generally form the active chorus for the spirit's songs. Those further away may contribute questions to the spirits, but they also pass comments and commentary about the performance among themselves in the darkness. Who sits with whom and how near they sit to the medium has some effect on the way the conversation goes. This audience repositions itself over time as people shift around in the darkness, sometimes moving closer to the medium if some topic of interest to themselves comes up or changing position to be near new conversation partners. Given the importance of the audience in co-construction of the performance, these movements and changes of grouping seemed important to track. I could diagram the initial positions of the gathered people before the séance began, but once the fires were put out and the séance was under way, the movements of audience members in the pitch blackness were difficult to follow. Only an infrared video camera would have made this possible, a technology not available to me at the time.

A major technical issue for my recording was the correct positioning of the microphone. For me, the problem was catching both the dialogic interaction between audience members and spirit voices next to the medium and also the conversations among audience members in the background—which often contained much useful commentary and critical observation that affected the performance process (Schieffelin 1995). In order to get all of this, I usually positioned myself in the second tier of the grouping around the medium—close enough to hear even the most minute of spirit voices, but also able to pick up conversations going on at the periphery behind me. From this position I could often pick up spirit speech that people further away could not hear, but also record the important audience background conversation while the spirits and chorus were engaged in the midst of song. In this way I was able to identify the role of multiple, differently focused points of awareness in the development of the performance as a whole.

Transcription

Once a recording is completed, transcription can begin. The first issue to confront in listening to my tapes was what to include in the written-down text and what to leave out. My experience in recording had already suggested that I could leave out virtually nothing. Coughs, laughter, cat-calls, dog-fights, small noises of the night, and even (on one occasion) the sound of a fart—all had the capacity to affect the way the performance went, and so had to be included. The most prodigious and time-consuming task was simply untangling what was being said by the individual voices among the multiply overlapping conversations and spirit songs in the longhouse. The whole was made more complex by the fact that the performance was entirely held in a Papuan language that I knew well but whose finer nuances I could not always penetrate despite nearly five years of field experience.¹ Translation had to be undertaken as transcription went along.

I always transcribed with three Bosavi informants seated around the tape-recorder to help me hear—two of whom had attended the séance, and one who had not. The first two could use their memory of the events and their native ear to help decipher the conversations in the Bosavi language, while the third provided a fresh ear and outsider's perspective. We had to

¹ Many field workers employ assistants from among the indigenous people to transcribe their tapes for them. Bosavis at the time were non-literate, so I had to transcribe the tapes myself.

cover the tape inch by inch often with multiple listenings and considerable discussion about the words that we were hearing as well as how they were meant. As a piece of work calculated simply in terms of brute force, transcribing these séances was a real bear.

But the brute work of being able to hear, let alone separate, the conversations for writing down was only one problem facing this transcription process. Another, potentially more difficult question was the level of performative detail it was necessary to transcribe. It is one thing to simply transcribe the words—it is another to transcribe them performatively. This is of particular importance for verbal art performances such as oral poetry, where meter and rhythm, changes in pitch and tone, speed of delivery, and other speech effects are of the essence and must be encoded in the transcript—but, in principle, speech effects are important in any performance with a verbal component. This was certainly true for Bosavi spirit séances. There was a good deal of emotional expression among the audience members throughout the performance, as well as verbal performance effects from the medium: each spirit that came up had a different and distinctive voice, and audience members sometimes could identify it by its speech before it had given its name.

In principle, transcription aimed at recovering this kind of linguistic detail in a meticulous manner entails splitting the performative activity into various levels of analytically distinguishable speech effects and transcribing or annotating them separately in parallel with the lexical text. On the linguistic level alone, there are numerous paralinguistic features such as intonation, pitch, loudness, rapidity of delivery, and so on that may be important in the performance. Documenting this accurately in a transcript involves an immense amount of work and training. I managed to accurately transcribe the points of conversational overlap and the background speech that accompanied the songs, but the encoding of all the finer details of speech production, although I could hear them, was simply beyond me.

Yet it was crucial that the effect of these performative elements of the language not be lost. I decided upon an ethnographic approach to resolving the problem. My fundamental aim in making a transcription of a spirit séance was to understand how Bosavi spirit mediums and their audiences co-constructed the imaginative reality of the spirit séance and worked within it to performatively pursue particular agendas. What were the genre conventions and what strategic moves were available to the players within them? How did audience and medium maneuver their agendas through the developing performance? These questions meant that it was more important for me to note that a speech effect could be identified at a particular place, and that it had a particular significance to those present (and for the

performance) than it was for me to transcribe it with full paralinguistic accuracy and detail in the text. So I turned from meticulously transcribing performative paralinguistic to detailed annotation of indigenous perception of performance effects.²

I began asking my assistants what they perceived was going on at points where either I or they noted interesting inflexions of speech or emotional expression. I wanted to know what people were reacting to, why they reacted that way, and what the outcome was in the development of the séance. I found that my informants were able to give me a knowledgeable running commentary and evaluation of the performance all the way through. They knew what was going on beneath the words, what people seemed to be up to, and they could give the reasons for their perceptions. In addition, they provided canny reflections on the performance based on memory and hindsight as well as material on the tape: noting what they had thought was going on at the time and what they thought about it now and why. They also brought to the discussion a well-informed knowledge of the séance genre. I recorded their comments and interpretations in extensive annotations to the transcript in the places where they occurred. Where there was disagreement I recorded the difference of opinion or detailed the course of discussion if they reached a resolution.

What emerged from this approach was an enormous wealth of séance lore, insightful perception, thoughtful interpretation, and performative insight from experienced séance attenders, which could explain the speech effects at various places in the transcript, but in no way could have been deduced from a meticulous transcription of paralinguistic. The performance did not reside in the “objective facts” of the event but in the working understandings, strategic moves, and developing situation of its participants—much of which developed at a level above, or at least outside, the level of the words of the transcript. We will return to this problematic point.

² This approach does entail a loss and would not be sufficient for those scholars who wish to study the details of the linguistic means by which particular speech effects and performative effects are brought about. In this transcription I had to restrict myself mainly to identifying and annotating the existence of performative effects and understanding their significance and consequences.

Context of the Performance

The context of a performance is one of the important aspects that makes it what it is. I do not mean here the theater or longhouse, nor, at this point in the discussion, do I mean *Tradition*: the genre values, conventions, and expectations that people bring to these performances (although these are very important). I mean rather the social, historical, and ethnographic features that place the performance in its local ethnographic and historical context and are significant to its meaning. It is, for example, necessary to know something of the particular social and political situation of Athens in the fifth century B.C.E. in order (nowadays) to understand why audiences of Aristophanes thought his plays were funny.

Bosavi séances were largely concerned with curing the sick and identifying witches—inevitably issues of considerable local importance, which frequently had significant political repercussions throughout the local communities. Séances occurred at particular junctures and in specific contexts of local circumstances. What went on in séances often played a strategic part in local social and political events. Indeed, in Bosavi at least, many of these performances did not make full sense outside the ongoing developments of the local situation. An important part of the meaning and role of these performances was only to be found by investigating the circumstances that surrounded their being staged. Indeed, a great deal about the unspoken undercurrents of the local social and political situation arose in the performance itself, and could be tracked by annotating the transcript about the allusions and references that arose in spirit songs, and why particular participants raised particular issues. I often had to elicit a vast amount of local history and sociology in order to fully appreciate how and in what way the performance did its work, or why it had the effects on particular audience members that it did. As noted earlier, a great deal of it was elicited for annotation of the transcript during the transcription process itself.

At the same time, and apart from social and political issues, séances were also sources of enjoyable entertainment for Bosavis and subject to aesthetic evaluation by villagers who had developed considerable sophistication in judging the fine points of these performances. I elicited considerable material from informants about these matters, to the point where it was possible to develop an outline of Bosavi criteria of aesthetic judgment and gain an insight into their understanding and appreciation of these performances. This, of course, was an important goal of performing an ethnographic transcription in the first place.

The completion of a fully translated and annotated transcript of a Bosavi séance was, as I have indicated, an immense amount of work. On average it required about 100 hours to transcribe and annotate each hour of performance. Although transcription is always a tedious process, I would think that this amount of time is unusual as such things go. Part of the reason was that I could not give the task over to a research assistant. But I also chose to transcribe it myself because I wanted to develop a deep familiarity with the material. In the end, had I done otherwise, I would have missed most of what turned out to be the really interesting performative content. The annotations of the transcript were as valuable as the transcript itself, and even a detailed paralinguistic transcription would not, of itself, have picked up what the Bosavis found significant and interesting—it would only have signaled that there was perhaps something interesting there to investigate.

Pitfalls of Detailed Transcription

Once the transcript is completed, one of the attractions that it holds for the researcher, if not the general reader, is that it seems to provide a highly detailed and meticulous record of what went on in the performance that can be used to revisit and minutely scrutinize its terrain and discover and reflect upon the means of its inner workings. As such the transcript can be an extraordinary tool for the close analysis of the performative process. While a good transcript can afford this opportunity, trusting the transcription has some serious pitfalls that have a lot to do with accuracy and detail. Although there are a number of problems here, I will restrict myself to discussing only three, all of which have to do with the dangers of reading too much into the transcript.

One we have already touched on is the problem of unilluminating details—that is, when details of the transcript indicate that something is potentially there but not what it is. The main example is places where the transcript indicates points of paralinguistic intensity or transition, but not what they signify. It is only through additional work with an informant that it is possible to determine whether such conversational markers are perceived by the speakers themselves, whether they have performative significance, and if so, what that significance is. This is not revealed in meticulous linguistic transcription by itself.

A second problem is that a carefully detailed and documented transcription often reveals events or complexities in a performance that were not apparent to the participants (or all of them) at the time. In my recordings of Bosavi séances, there were several instances where verbal material

appeared on the tape that nobody seems to have heard during the performance. There were also many instances where linguistic materials were heard by only a restricted number of people during the performance, or were heard but misunderstood. In addition, my transcription informants sometimes saw significances in the material in hindsight that had not occurred to them during the performance itself. Such insights by hindsight can potentially throw a new light on what was going on in the performance, but do they really replace what people thought was happening at the time? Many questions arise here. Can something that was patently not perceived during the performance be said to have had performative effect? What part does performative material that was mistakenly understood at the time play in the way the performance actually unfolded?

So it is necessary to observe caution in working out performance dynamics based on details uncovered in the transcript that participants were not aware of while the performance was going on. In the end, resolving these issues comes down largely to a careful judgment based on a combination of informant discussion and transcript. At the same time, the transcription of unheard detail suggests an interesting line of research, not well pursued in performance studies, namely the study of how a performance proceeds through the imperfections and miscommunications of its interactions with its audience. A detailed and well annotated transcript will have something to contribute to this.

Finally, one of the most difficult aspects of performativity to recover in a transcript is the pace, or dramatic timing of events in relation to one another, and this affects the points at which one performative move may be perceptually distinguished from another. The difficulty of transcribing timing opens the way for another potential problem made possible by a detailed transcript. I call this the fallacy of misplaced detail—or the Rodney King problem. Rodney King, it will be recalled, is a black American man who was picked up by the Los Angeles police for acting suspiciously and was found to be drunk. He was beaten mercilessly for a prolonged period by the police for, as they said, resisting arrest. Fortuitously for King, the beating was recorded on videotape by a bystander and became the basis for a controversial and celebrated trial in which King sued the police for racially motivated brutality and use of excessive force. The tape shows a beating savagely delivered by several policemen to a man who at first briefly seems to resist, and finally appears to be trying to protect himself with his arms as he is beaten mercilessly to (and on) the ground.

A simple viewing of the tape (which is less than five minutes in length) would seem to leave the viewer in no doubt that it is simply a brutal beating. The defense, however, took the jury through the tape frame by

frame, showing that the way that King's raised arms and movements during the beating could be interpreted at each point where he received another blow as representing a vigorous and dangerous move, which was appropriately interpreted at each point by the police as a continued aggressive resistance justifying a continued use of force. The defense won the case in what many observers concluded was a serious miscarriage of justice. Here, the minute, piece-by-piece breakdown and analysis of the event contradicted what was plainly visible to everyone who saw the brief and confusing action on the video as a whole. The defense succeeded in persuading the jury that the event consisted of, or could legitimately be viewed as, a large number of small incidents linked together consecutively rather than as a single performative flow. With the flow removed, the whole was reduced effectively to less than the sum of its parts. Or worse: some would argue that the evidence was actually falsified by this way of breaking it down and presenting it.

This kind of problem can also bedevil the interpretation of a detailed transcription of any performance no less than it did the Rodney King tape. What one can interpret from minute consideration of details teased out of the transcript—but lost to the participants in the flow of events at the time—does not by itself necessarily provide a basis for a better understanding of what was “really” going on. It can simply be mistaken.

A Note on Using Transcriptions

The foregoing suggests that the transformation of performance events into an objectified textual format brings with it a particular impoverishment of the material. This is not because a transcript can never contain everything. On the contrary, impoverishment can result from excess. Rather it is because, in the end, there are a number of qualities intrinsic to the flow of performance that cannot really be transformed into text, or rather, which become significantly altered when that is attempted. This is particularly true with the dynamic, evanescent qualities of performance—pace and rhythm, movement and trajectory, tension, climax, and release. While appropriate marks are placed to indicate and qualify various aspects of these features in the transcript, the sense of the flow of performance itself, with its qualities of emergence and participation, is lost. The more transcription strains to encompass those effects, the more it must engage complex analytical processes if it is to record them in text at all. As we have seen, the transcription breaks the flow of moving events into a series of constitutive elements, themselves split into several registers of perception according to

different domains of analysis (linguistic, musical, kinesthetic, and so on) each of which may entail several levels of annotation. Paradoxically, then, as transcription tries to become more objective, neutral, and transparent to its material, it is forced to become more analytic and interpretative.

In my view, it is useless to worry about these problems. The issue is how to creatively get around them, and this cannot best be done using the transcript as text alone. I suggest that a transcript is most useful for researching performance when it is read alongside of, or in conjunction with, the recording of the performance from which it was made, allowing each to enhance and play off each other. Here the transcript is used to track and clarify the detail of what is going on, while the tape recapitulates and renews the sense of performative flow. It is through the emergent play between the transcript and recording used together that work with a transcript of performance becomes most useful and revealing.

*University College London
University of London*

References

- Schieffelin 1995 Edward L. Schieffelin. "On Failure and Performance." In *The Performance of Healing*. Ed. by Carol Laderman and Marina Roseman. London: Routledge. pp. 59-90.
- Schieffelin 1998 _____. "Problematizing Performance." In *Ritual, Performance, Media*. Association of Social Anthropologists, 35. Ed. by Felicia Hughes-Freeland. London: Routledge. pp. 194-207.