

Captured on tape: professional hearing and competing entextualizations in the criminal justice system*

MARY BUCHOLTZ

Abstract

A growing body of discourse-analytic studies demonstrates that within the legal system, spoken language that undergoes entextualization is transformed in a variety of sociopolitically consequential ways. Through the analysis of a legal case involving the institutional entextualization of incriminating language—an FBI summary log of wiretapped telephone calls between suspected drug dealers—the article argues that practices of professional hearing and transcription on behalf of institutions of law enforcement systematically work to the disadvantage of suspects and defendants. At the same time, discourse analysts who intervene in the justice system as linguistic experts offer competing representations of suspects’ and defendants’ language rooted in their own practices of professional hearing and their own institutional goals. Legal entextualizations, whether weighted in favor of or against defendants, present dialogically constructed representations of discourse as monologic and objective facts in order to foster an ideology of institutional neutrality. This process allows legal institutions to “create their own reality” by regimenting talk through textual representation in order to promote their own authority and interests.

Keywords: transcription; language and the law; entextualization; power; representation; reflexivity.

1. Introduction: professional hearing and the law

Within the criminal justice system, the institutional regimentation of often unruly talk is constituted in important part through the phenomenon of *professional hearing* (see also Ashmore et al. 2004). Modeled on Goodwin’s (1994) closely related notion of professional vision, the concept of professional hearing calls attention to the ways in which institutions

foster distinctive ways of making sense of language. Like the archaeological practices that Goodwin documents to illustrate professional vision, professional hearing is part of what it means to participate in the activities of an institution; linguistics, for example, has a well-developed set of techniques of professional hearing, which allow analysts to notice and examine aspects of language that the lay listener may overlook or misinterpret. While professional practices of seeing and hearing are crucial in enabling institutional activities to proceed, they also shape how those activities unfold. In his analysis of the trial of the white police officers whose beating of black motorist Rodney King in Los Angeles was captured on videotape, Goodwin uses the concept of professional vision to account for discrepancies between the police officers' interpretation of this videotape and the interpretation of the lay public. Similarly, a number of forensic linguists have shown that through institutionally valorized practices of professional vision and hearing in the criminal justice system, contestable interpretations of events may gain the official status of objective and hence undisputed truth.

An area of ongoing interest in this regard is the role of linguistic representation in transforming spoken language into a more tractable written text. In the legal domain, professional hearing is often mediated in some way through documents such as transcripts and other written representations of spoken language. Until recently, texts of this kind were typically the sole official source of evidence for criminal wrongdoing originally documented via recordings, and they continue to be used as evidence in many cases. Such representations of speech thus have a particularly powerful role to play in how professional hearing shapes legal outcomes. Much of the research on transcription in the justice system focuses on how legal professionals—from the police officer who makes an initial arrest to the court reporter who records the courtroom proceedings—engage in practices of professional hearing. In this article, I likewise consider the practices of such professionals, but I also extend my analysis to include my own practices of professional hearing as a linguist.

Drawing on data from a criminal case that relied on the written representation of speech and for which I served as a pro bono consultant, I argue that the professional hearing of those who produce such documents for use in the legal domain is not simply a neutral technical skill but a professional practice through which differing institutional perspectives, interests, and versions of reality may be pitted against one another. In the case I examine below, differences between the law enforcement agency's transcript and my own version systematically worked in favor of the interpretations of the officials who recorded the original evidence. This transcriptional process may be understood as coercive in that it puts

words into defendants' mouths against their wills and their best interests. Yet linguistic researchers who enter the legal realm as consultants or expert witnesses may in turn undermine the authority of such official transcripts by bringing their own practices of professional hearing and institutional agendas to the task of entextualization. Thus competing institutional representations of the same event often become contests over which representation will enter the record as the "official version," with the outcome consistently weighted in favor of the representation structurally endowed with greater authority in a given institutional arena.

2. Getting a hearing: entextualization in listening and the law

Transcription is a form of entextualization, a process whereby language becomes detachable from its original context of production and is thus reified as text, a highly portable linguistic object (Bauman and Briggs 1990; Silverstein and Urban 1996). In transcription, embodied language use is extracted from the richly contextualized social interaction in which it is originally produced and through its recontextualization in writing is rendered static and fixed on the printed page. The original dialogicality of speech is replaced by the monologic authority of the transcriber, who determines what was said and how to represent it. This process of "creating one's own reality" via linguistic representation is as central to scholarly transcription practice as it is to other forms of institutional transcription (e.g., Blommaert and Slembrouck 2000; Bucholtz 2000, 2007; Jaffe in this issue; Mishler 1991; Mondada 2002; Ochs 1979; Parker 1997; Preston 1985; Rosenthal in this issue).

The entextualization of spoken language through transcription is a semiotic practice that is both material and ideological, in that it involves the auditory apperception of the speech stream, the interpretation of what is said, and the representation of that interpretation in written form. These processes are agentive rather than mechanical, and they often rely on the professional beliefs, expectations, norms, and interests of both the auditor/interpreter/transcriber and the institution she or he represents. Such ideologies and interests, however, remain largely unavailable to scrutiny in the everyday workings of the institution (see also Park in this issue). In fact, the invisibility of ideology in much of professional practice is itself ideological, in that objectivity and neutrality are central ideological pillars of most professions, from the justice system to linguistics. These ideals require the marginalization and delegitimation of some practices as nonobjective and non-neutral, and in so doing, they position objectivity and neutrality as the normal, typical result of

professional practice. Yet as scholars have abundantly shown, the notion of objectivity is in fact a sociohistorically specific form of professional vision that can never be fully realized in practice (Daston and Galison 2007). Hence even institutional agents who adhere strictly to the highest professional standards of objectivity and neutrality inevitably bring their profession's point of view to their interpretations and representations of what they see and hear (for this phenomenon in discourse analysis, see Blommaert 1997).

All entextualizations are necessarily arrived at dialogically and are thus inherently "double-voiced discourse" (Bakhtin 1984: 185). By erasing the intertextual gap between the originary and the entextualizing discourse (Briggs and Bauman 1992), however, institutions are able to portray their entextualizations as objective and transparent, thus fostering an ideology of institutions as neutral and disinterested entities. Through this process, entextualization underwrites institutional authority (Park and Bucholtz in this issue). In the legal realm, the acts of official hearing and recording authorized by the institution may thus be persistently stacked against those who lack institutional power (i.e., usually the defendants) even without deliberate intent to tip the balance in favor of the institutionally powerful.

The power that underlies professional hearing is conferred in the first instance through the right to listen in. Within the surveillance culture of postindustrial society, governmental spying and eavesdropping on citizens has become commonplace (e.g., Ball and Webster 2003; Lyon 1994, 2001; Staples 1997, 2000). Current surveillance techniques not only realize but improve upon Jeremy Bentham's eighteenth-century dream of the penal system as panopticon (Foucault 1977), for in addition to scrutinizing the behavior of potential, suspected, and convicted criminals with the all-seeing eye of authority, today's institutions of justice also monitor the talk of such speakers with the all-hearing ear of recording equipment: the *panaudicon* (cf. Rice [2003] for a somewhat different use of this term in the context of institutional power). In this way, suspicious speakers' private talk enters into public circulation at two different moments: first as it is exposed to the official representative of the eavesdropping institution and then as it is held up by the prosecuting institution as incriminating evidence.

Thanks to rapid improvements in technology, audio and video recording has become a common tool for gathering legal evidence. Indeed, in a few notorious cases of police brutality this ready access to recording equipment has allowed members of the public to turn technology against abusive and out-of-control officers (e.g., the beatings of Rodney King in Los Angeles in 1991 and of Donovan Jackson, an unarmed black teen-

ager in Inglewood, California, in 2002); in these situations and others, surveillance, scrutiny from above, is challenged by *sousveillance*, scrutiny from below (Mann et al. 2003). However, in most cases in which recordings are used as evidence—and these are still relatively rare—law enforcement officials as institutional representatives control both the means of recording criminal evidence and the products of such surveillance events, such as interrogations and confessions. Not only are these recordings typically unavailable to the public, except perhaps in highly selective and edited form, but they may even be unavailable to the juries who depend on such forms of evidence to determine a defendant's guilt or innocence.

Given the law's longstanding privileging of written texts over spoken language (Tiersma 1999; Walsh 1995), as well as the additional costs and complications of introducing technology into courtrooms, until recently, often only the transcript was admitted into evidence and thus was the only source of information that jurors could examine in coming to a verdict. Such texts are far from neutral records of what was said. Studies of the entextualization of police interrogations and court proceedings report a consistent pattern of bias in legal transcription in favor of those who hold the greatest institutional power—judges, lawyers, and police officers—by standardizing the language of powerful speakers or representing them as sympathetic figures (e.g., Coulthard 1996; Eades 1996; Walker 1986, 1990). Conversely, transcripts produced in the course of the judicial process may misrepresent the speech of those suspected or accused of crimes, often in ways that are consequential to the outcome of trials; linguistic experts have used their professional hearing of language, dialect, and speech style features to demonstrate the inaccuracies of such legal transcripts (e.g., Blackwell 1996; Bucholtz 1995; Esau 1982; Patrick and Buell 2000). Wide linguistic gaps between official representatives of the legal system and those who become enmeshed in its workings can be particularly problematic, for in such situations the representational process is often further mediated by the presence of an interpreter (Fowler 2003; Jacquemet in this issue). In light of the overwhelming evidence that forensic linguists have marshaled regarding the limitations and flaws of legal transcripts, a recent book on linguistic issues in the law, aimed at legal professionals, strongly recommends that police officers routinely audiotape and videotape their encounters with suspects, including interrogations, and that transcripts only be used by juries as an aid to understanding recordings, not as evidence in their own right (Solan and Tiersma 2005), a practice that is now becoming standard in most courts.

Even when electronic recordings of speech are admitted into evidence, however, these may have problems of their own. Recording technologies remain far from perfect: video images are often dark, blurry, or grainy,

and the quality of audio recording may be compromised by static, background noise, or the low intensity of the speech signal. The resulting ambiguity of recorded evidence is one reason that institutional techniques of interpretation and meaning making—professional hearing and seeing—often mediate the lay public’s access to surveillance recordings.¹ Thus the entextualization process is not avoided, only altered, when speech is recontextualized as an audio or video recording rather than as a transcript. Indeed, research shows that every aspect of entextualization within the legal system subjects language to high-stakes recontextualization, from manipulation of the recording process itself (Shuy 2005: 159–164) to coercive interrogation techniques (Gibbons 1996) to the negotiation of what should be included in a transcript as court proceedings unfold (Walsh 1999) to the oral performance of a transcript in a courtroom argument (Cotterill 2002). Even the seemingly simple matter of what is said in a recording has been shown to be open to dispute (Bowe and Storey 1995; Fraser 2003). Such problems, which are inherent in the use of recorded evidence, are obscured when, as in the case discussed below, the written entextualization of recorded speech is the primary source of evidence in the legal setting.

Moreover, language scholars are not immune to the influence of professional hearing in their own transcripts. When linguistic experts enter the legal arena, we may be tempted to romanticize ourselves as heroic underdogs advocating justice and speaking linguistic truth to power, but the evidence of our own practices may be rather more complex. In the courtroom, the professional hearing of language experts and the resulting entextualizations of speech that it produces are not unmediated truth but one contestable form of institutional authority and representational practice among others.

3. The institutional construction of suspicious speech

The data analyzed below are taken from a case in which I was involved as a pro bono consultant based on wiretapped recordings of incriminating telephone conversations. I am not a specialist in forensic linguistics, and in this case I was approached by the public defender as a last resort due to my research specialization in African American English. The attorney hoped that a linguist might be able to find something in the recorded evidence that would undermine the case against her client. In particular, since the defendant was African American, the California-based public defender, perhaps influenced by the flurry of public attention to differences between Standard English and African American English following

the 1997 Oakland School Board Ebonics controversy (e.g., Ramirez et al. 2005), suggested that there might be other, more benign interpretations of the apparently damning recorded speech if linguistic and cultural differences were taken into account. The defense attorney did not have any initial concerns about the transcripts, which she viewed as straightforward representations of the recordings. However, the most obvious and pressing linguistic issue that emerged from my review of the recordings was not one regarding dialect differences but rather the mismatch between what I heard on the tapes and what I saw in the written record.

Several ethical issues in this project must be addressed at the outset. First, it is important to note that given my role as a consultant for this public defender, my task was to apply my linguistic knowledge in ways that would help her to defend her client—whether by analyzing phenomena that required linguistic expertise, by identifying issues that the attorney might have overlooked, or by demonstrating that the linguistic case she hoped to build lacked adequate evidence. In doing so, I sought to combine my professional commitment as a linguist to analyze data empirically with my agreement to assist the defense to the extent that I could; if I had discovered something in the recordings that was beneficial to the prosecution, my obligation was to share that information with the defense, not to alert the prosecuting attorney. Hence I was quite explicitly not a disinterested researcher and my professional hearing was calibrated to interpret the recordings in a particular way. I did not ever deliberately misrepresent what I observed, but I was fully aware that I consistently examined the recorded material with a view to whether an innocent interpretation was possible. My approach was in keeping with the “reasonable doubt” standard that juries are charged to uphold, but it also conformed with my usual professional hearing practice as a discourse analyst: not to rush to assign the most immediately obvious and apparent interpretation to data but to weigh all plausible possibilities.

A second ethical concern is that by playing the data recordings in public or offering transcripts of them in published form, the research becomes, in effect, an additional level of surveillance, eavesdropping on people who are unknown to the researcher and to the study’s readership and who are powerless to prevent such surveillance. To be sure, this is not eavesdropping in a legal sense, in that it is the government, not the researcher, who carried out the recording and made it publicly available; moreover, these actions were performed in accordance with the law. Nonetheless, scholars are not off the ethical hook simply because their data were legally obtained—the barest minimum ethical standard—for the use of private conversational data without the speakers’ consent raises serious concerns. Although there is some indication that the speakers in

the wiretapped telephone calls were aware that their phone might be tapped, no doubt they would have presented themselves very differently if they had known that an audience of academics would listen to or read transcripts of these recordings. Nor to my knowledge were the clients ever even informed that I had been given access to the data and was involved in their case; I never met them or had any other contact with them. It is in fact a common practice for public defenders to enlist expert assistance without consulting their clients or introducing them to those experts, in order to prevent the expert from being accused of bias.² But as Cunningham and McElhinny (1995) note, this practice further disempowers an already extremely vulnerable population. Despite these concerns, I feel it is important to present this material for what it reveals about the unscrutinized power of official transcripts to shape the outcome of legal proceedings.

3.1. *Listening in: entextualizations of wiretapped recordings*

The data I present involve wiretapped phone calls of a group of suspected drug dealers in California's Central Valley. In May 2000 I was contacted by the federal defender regarding an upcoming hearing to extend the Federal Bureau of Investigation's authority to wiretap the suspects' telephone. The FBI justified its request based on the use of what its agents claimed was coded language. In the words of one agent involved in the case in a written statement to the court, "Due [to] the background of this case, the deliberately vague and guarded language used by the parties to the conversations, and the overall context of the calls, I believe the calls concerned the sale of cocaine" The federal defender hoped that I might be able to develop an argument that the terms the speakers used in the phone calls were not in fact a code and that the words in question also had non-nefarious meanings, possibly from slang or African American Vernacular English. I provided the attorney with the best evidence available from the recordings for such a position, but I also informed her that in my opinion the greatest contribution I could make was to call into question the FBI's written representations of the suspects' conversations.

It is clear from the FBI agent's statement above that an orientation to possible surveillance is itself suspect in a surveillance society. As many pro-surveillance commentators have remarked in their rejoinders to public concerns over the erosion of civil liberties in the post-9/11 United States, "If you have nothing to hide, you have nothing to worry about." Thus the suspects' awareness that their phone calls might be monitored and the measures they took to foil possible eavesdroppers were used as

evidence against them. Similarly, in discussing a criminal case involving surreptitious recordings made by FBI agents, Shuy notes that the mere existence of such recordings can foster a presumption of guilt:

In court cases which involve tape recorded evidence, there is a public predisposition toward guilt. The reasoning is emotional, not intellectual. It runs something like this: "If the FBI suspected that these men are guilty and succeeded in getting tape recorded evidence to that end, the defendants are probably guilty. . . ." No juror would admit this, of course, but the emotional set is there. (Shuy 1986: 234)

Such issues are especially important to keep in mind in a case like this one, where the evidence may appear to point clearly to criminal activity. It is precisely in such cases, I would argue, that protections of individual rights must be most firmly upheld.

The wiretapping case hinged on several phone conversations that the FBI asserted were indicative of drug-related activity. The federal defender sent me the audiotapes of these examples along with the FBI's entextualization of them. In this case, the wiretap recordings themselves were not admitted directly into evidence and no transcript was created. Instead, the FBI submitted to the judge presiding over the hearing its own logs of the recordings at issue, which were made by the FBI agents who monitored the tapped phone line. The logs are not word-for-word entextualizations of the speech on the tape recordings but brief summaries, which necessarily involved a great deal of interpretation on the part of the FBI agents. It is not unusual to base criminal cases or hearings on evidence created by the law enforcement agency itself; challenges to the agency's entextualized representations are rarely mounted by defense attorneys and are even more rarely successful.

In the examples below, the FBI logs are reproduced exactly as they appear in the court documents except that names are changed to initials. My own transcripts are based on the tapes, which are of extremely poor quality. I was very conservative in what I felt I could confidently transcribe from these tapes, a cautious approach that admittedly was to the benefit of the clients in that it could make available less evidence against them. The goal of my analysis was to examine whether the FBI's interpretation of the suspicious utterances was the only or most plausible interpretation available. The first two examples below are those for which I was unable to propose an alternative hearing. The speakers in both examples appear to be metalinguistically negotiating a pre-established code (see also Shuy 1997), and in example (2) they do so at some length. I could offer no plausible argument that these examples should not be admitted into evidence and did not address them in the brief that was submitted to the court. The following analysis does not replicate the brief

itself but lays out the reasoning that informed the argument presented to the court regarding these entextualizations. (See the appendix for transcription conventions.)

(1) a. FBI log of telephone call between suspected drug dealers

[BP] CALLS [AT]. [T] TOLD [P] TO TAKE A DOVE AND PUT THE REST IN HIS ROOM. ALSO [P] SAID MEXICAN DROPPED OFF THE THING.

b. Researcher transcript of telephone call between suspected drug dealers

- 1 A: ##### May I help you?
 2 B: Yea:h.
 3 (We want a piece of) your loot,
 4 man.
 5 A: How much is it?
 6 B: A hundred.
 7 A: A hundred?
 8 B: Yeah.
 9 A: Just uh take a dove out of the stuff and put it on the ##
 (all right?)
 10 Take a hundred off the—
 11 I mean not a hundr- [@
 12 B: [All right all right all right!
 13 <sings/chants #####> @@@@
 14 A: Take a dove off # and put the rest in the pocket in my
 room.
 15 B: No problem.

Here the FBI singles out the word *dove*, which is in fact a common slang term for a particular quantity of crack cocaine (a piece of information I did not volunteer in my brief, but one which was already known to the FBI). It is quite clear in lines 10 through 13 that the speakers are orienting to lexical choice, as signaled by several features of the interaction. Thus in line 10, A initiates self-repair on his use of the term *a hundred*. He is interrupted by B, who shouts “All right!” repeatedly and then begins to loudly chant or sing (lines 12 and 13). B’s turn could be interpreted as designed to distract from or drown out A’s repaired speech (although the difficulty of understanding the content of the singing or chanting makes it hard to offer a definitive interpretation). In any case, when A successfully repairs his utterance to *a dove* in line 14, B accepts the rephrasing without further comment.

Example (2) provides an even more obvious instance of speaker orientation to coded language:

- (2) a. FBI log of telephone call between suspected drug dealers

WITH THE CAVE [apparent typographical error for *cake*].
HOW MANY PEANUTS WWERE [sic] THERE IN THE
PACKAGE FOR [B]? [T] LAUGHING IN BACKGROUND
ABOUT CODE. REGARDING [P]'S TRIP TO SEE [S] ON
SATURDAY.

- b. Researcher transcript of telephone call between suspected drug dealers

- 1 B: What's up,
2 dog?
3 A: What's up,
4 brother?
5 B: How you doing?
6 A: All right.
7 B: Hey.
8 You go- you y- got that cake,
9 right?
10 A: Huh?
11 B: You got that cake l- like I tell you,
12 right?
13 A: My what?
14 B: The cake,
15 goddamnit.
16 The red velvet cake.
17 A: Yeah,
18 the cake,
19 motherfucker.
20 B: And that peanut,
21 right?
22 The peanuts?
23 How many peanuts in that bag?
24 A: Huh?
25 B: How many peanuts in that bag?
26 A: How many ^peanuts?
27 B: Are in the bag.
28 A: There ain't no peanuts.

- 29 B: The peanut bag for (A),
 30 motherfucker.
 31 T: <in background> (Listen to) what he's saying,
 32 niggah.
 33 A: O:h.
 34 Oh-
 35 Oh!
 36 T: #####[1## 1]
 37 A: [1@@@1]
 38 T: #####[2## 2]
 39 A: [2@@@@@2]
 40 T: fuck you doing?
 41 A: .h u:h.
 42 T: @@@@^Oh g@od.
 43 A: There was two,
 44 baby.
 45 T: .h @@@
 46 A: Two.
 47 B: There was two?
 48 A: Yeah.

Here the FBI log takes a step away from the relatively objectively framed description of example (1a) to offer an interpretation of this exchange: speaker T is “LAUGHING IN BACKGROUND ABOUT CODE.” Although to insert one’s interpretation into the very data under analysis is considered bad form from a researcherly standpoint, it is difficult to argue with the FBI agent’s conclusions. *Cake* is another slang term for cocaine (*red velvet cake* seems to be a local innovation), and while I am unaware of the use of *peanut* in the same way, the speakers’ extended negotiation of the meaning of both terms suggests that the two words belong to the same semantic field. B introduces the term *cake* in line 8 with several markers of hesitation, which may indicate that the term is not one he usually uses; however, speakers hesitate for many reasons, and it is difficult to attribute motivation in this case.

Stronger evidence that B’s lexical choices are not commonplace language for these speakers is A’s persistent confusion about B’s meaning (lines 10, 13, 24, 26, 28), which he only overcomes when T urges him in line 31 to “(listen to) what he’s saying” (however, the quality of the tape makes the transcription of this line uncertain). In further support of this analysis, B continues to use the terms *cake* and *peanut* (with increasing emphasis and affective force) rather than rephrase his questions, despite A’s evident confusion. In lines 33 through 35, A produces a series of

change-of-state tokens (Heritage 1984), which seem to signal his dawning understanding of B's meaning, and after extended laughter by A and T and unintelligible talk by T, A answers B's question to the latter's satisfaction (line 46). Once again, however, the unintelligibility of T's utterances at a crucial point in the interaction (lines 36–38) makes it impossible to assert with confidence that the above analysis is either correct or complete. Moreover, while the use of these two or three drug-related slang terms may be indicative of criminal activity, the use of "vague and guarded language" in itself, which the FBI took to be strong evidence of wrongdoing, cannot be so easily interpreted as such.

If examples (1) and (2) are the FBI's strongest evidence against the suspects, (3) is a bit more ambiguous.

- (3) a. FBI log of telephone call between suspected drug dealer and presumed customer

[D] ASKED [T] IF HE HAD ANOTHER LOAD. [T] REPLIED YES. [D] WILL PICK UP LOAD AT [T]'S RESIDENCE PRIOR TO 9:30 A.M. [T] HAS TO GO TO WORK AT 9:30 A.M.

- b. Researcher transcript of telephone call between suspected drug dealer and presumed customer

1 D: Hello?
 2 T: Hello.
 3 D: Is that T?
 4 T: Yeah.
 5 D: This this is D,
 6 T: Hi how you doing?
 7 D: O:kay,
 8 o:h,
 9 I need another loa:d/loa:n. @@@

The FBI log records the crucial final word in this example as *load*, which the agent interprets as referring to a supply of drugs. However, *load* can refer to a wide variety of non-illegal items. But even more significantly, the word in question is not clearly hearable as *load*. Another possible interpretation is *loan*, which is at least as plausible in this context. The laughter and possible cues of embarrassment in D's self-presentation in lines 7–9 (such as hesitation and prosodic lengthening) would make sense in either case, since both are requests for assistance.

Example (4) is an even clearer case of the FBI's agenda driving its interpretation of the phone conversations. In this example the FBI log

“translates” the vague referring expression *half a thing* into the much more precise form *half a kilo* (i.e., of cocaine):

(4) a. FBI log of telephone call between suspected drug dealers

CALLS FOR [M], CHECKS TO SEE WHAT HE PAID THIS MORNING. [M] SAYS TWENTY TWO. ‘NEGRO’ WANTS TO BUY HIS TRUCK FOR 1/2 KILO AND CASH.

b. Researcher transcript of telephone call between suspected drug dealers

1 M: Yeah uh uh uh oh Negro want my truck right?

2 T: Uh huh.

3 M: So uh uh I tell him eighty five,

4 yeah but now he gonna swing a a: a a a half a thing and and and a few hundred.

There is no obvious support for the FBI’s interpretation here; unlike slang terms like *dove* and *cake*, which have a definite reference, albeit one that is obscure to outsiders, pro-terms like *thing* by their very nature can stand in for a wide range of other words. And although M’s recycling of his turn in line 4 just before producing *half a thing* could be argued to indicate nervousness or uncertainty about which term to select, M shows similar patterns of recycling in lines 1 and 3, where the same explanation is less convincing. As it happens, the FBI did not actually pursue this line of argument regarding M’s possible nervousness in line 4, for the agent’s interpretation of the expression was accepted by the presiding judge without any need for explicit justification.

In the world of pro bono legal work with public defenders, things rarely end well for the defendant, given that public defenders are seldom assigned winnable cases and often seek out a linguistic expert only when no other avenue presents itself. In this case the judge rejected all of the arguments I laid out and ruled that the FBI’s logs provided sufficient evidence to support the extension of the wiretap warrant. Yet as I have argued, the logs did not simply report but actively constructed some of the linguistic evidence justifying this decision. When, as here, the entextualized result of professional hearing displaces the recordings on which it is based, it is difficult to reinsert the auditory material into the discourse context. Hence, institutionally sanctioned entextualizations come to be accepted as reflections rather than representations of prior speech events.

A larger issue that I did not pursue in the brief regards the different representational effects of word-for-word transcripts of speech versus

logs or other written documents that summarize the gist of what was said. That is, regardless of the accuracy of the FBI's logs, it is possible that such entextualizations may represent speakers in systematically different ways than do transcripts. For instance, in example (2) above, contextual details such as laughter that appear in the transcript version are reduced to a bare and incomplete description in the wiretap log, which strips the discourse of everything but potentially incriminating content; such contextual details may invite judges and juries to grant greater sympathy to speakers represented in transcripts than in logs. It is therefore likely that law enforcement officials prefer logs over transcripts not only because they are quicker to create and more concise but also because they can highlight incriminating aspects of the discourse while omitting details that may lead judges and juries to draw their own, perhaps undesirable, conclusions.³

Although both forms of representation seek to suppress the inherent dialogicality of discourse by presenting the speech of others as though unaffected by the entextualizing process, transcripts obscure the voice of the transcriber and privilege the voice of the speaker, while logs subordinate the voice of the speaker to the narrating voice of the institutional monitor. In Bakhtin's (1984) taxonomy of discourse, transcripts may be classified as "objectified" discourse, which purports to present the voice of the other in an unmediated way, while logs are "direct referentially oriented" discourse (1984: 187), in which the voice of the self is paramount. Both, however, are ideologically put forward by their institutional sponsors as "single-voiced discourse" (1984: 189) in that they do not acknowledge the role of the entextualizing agent in creating the final text. In this way, institutionally sanctioned textual representations of speech take on the status of neutral and objective fact and hence are rendered largely immune from challenges to their authority. Yet such documents may be far from neutral: condemnatory government entextualizations in effect perform the speech act of accusing, thus appropriating a role that is by law reserved for the prosecution.

Although the entextualizing authority of the institution typically goes unchallenged, it is nevertheless vulnerable to attack from rival institutions with competing entextualizations. Thus, in the case discussed above, I positioned myself as an academic authority with special expertise in speech and its representation and with institutional backing from my own academic institution and the office of the public defender. Both in the brief I submitted to the court and in my discussion of it in this article, I presented my own transcripts as more or less authoritative in order to raise concerns about the FBI's representation of suspects' speech in writing. Acting as an institutional representative in order to achieve particular

goals, I offered an analysis that shored up the legitimacy of my own entextualization. I did not problematize the process of transcription itself, nor did I highlight questions about my own transcription decisions. But to address the question of competing entextualizations in a more theoretically adequate way, the solution to problematic transcription is not simply to appeal to the original recording, since as researchers of language know, even experienced transcribers may hear different things on the same tape. Ashmore et al. (2004) note that such unreflexive privileging of the recording, or what they call “tape fetishism,” goes hand in hand with professional hearing, and they point out that academic transcribers are just as susceptible as other professional transcribers to hearings that serve institutional (including theoretical, methodological, and disciplinary) goals. Such issues call attention to the importance of reflexivity in our own research process as well as in our confrontations with the authority of other institutions (Bucholtz 2001).

The above case suggests that when different practices of professional hearing—here, legal versus linguistic—come into conflict, it is not simply a matter of institutional power versus disinterested scholarship. Rather, all entextualized representations are inevitably inscribed with the interests of those who produced them. In institutional contexts such as the justice system, the starkest difference is not between competing ways of hearing and transcribing but between experts—whether linguistic scholars or agents of the law—who are granted the authority to create entextualizations in the first place, and those whose speech is subject to recording, transcription, and interpretation and thus made to speak either in support of or in opposition to their interests. The central issue, then, is not misrepresentation per se (and thus not simply a question of transcribers’ responsibility to represent recordings accurately), but the way in which institutional contexts structure official records to control the lives of subjects.

It is important to acknowledge, too, that as the entextualizations discussed here enter new institutional contexts they may either gain or lose the luster of authority withheld or conferred in prior contexts. In particular, by decontextualizing my own transcripts and the FBI’s rival versions from the legal context for which they were originally created and recontextualizing them within the academic arena of scholarly publication, I suspect that I may have tipped the balance of institutional authority in my own favor. Although my attempt to promote the version of reality represented in my own entextualizations was defeated in the legal realm, in academia, where the institutional authority of my entextualizing practices is more secure, it may be found more persuasive. Yet any such victory is Pyrrhic at best, for it does not alter the ultimate conviction and imprisonment of the clients on whose behalf my arguments were made.

4. Conclusion

As spoken discourse enters into chains of entextualization within the legal system, it is institutionally transformed in a variety of ways: from speech to writing, from dialogue to monologue, from interaction to evidence. Professional hearing is the first step in this chain by determining what was said so that it can be inscribed into the public record. I have shown in this article that in FBI wiretap logs of the telephone conversations of suspected drug dealers, ambiguous speech is recontextualized in agents' narratives as damning evidence of criminal activity.

Thus in the representation of speech as legal evidence in criminal cases, questions of interpretation are powerfully shaped by the uses to which entextualizations are put. That is, by virtue of the practices of professional hearing that they acquire within the justice system, institutional representatives charged with the task of entextualization are able to coax or coerce a recording to say what they want it to say, even without being fully cognizant of this agentic act. As Mondada (2002) puts it, in legal transcription, it is such details that categorize a speaker as "guilty" or "innocent." The power of the written record thereby enhances the institution of the law and its authority to make its own reality, as well as the reality of those caught up in the legal system. Like the professional vision documented by Goodwin (1994) in the highly publicized trial against the police officers who beat Rodney King, the mechanisms of this institutional power and the questionable hearings that result shape the outcomes of the legal process in ways that systematically and dangerously disadvantage the speakers whose words are subject to professional representation. Within the institutional context of the legal system, the trajectory of discourse from speech to recording to transcript or other entextualizing document tends to erase its own tracks, leaving only the final entextualization as the authoritative representation of events.

As surveillance culture continues to authorize and normalize the scrutiny, representation, and circulation of subjects' speech by institutional agents, entextualizing practices such as those examined in this article will continue to create a version of reality that is in accordance with powerful interests. Such textual transformations of discourse into suspicious speech often stand uncontested as the institutionally authorized version. But in some cases, rival representations by academics and others challenge these officially sanctioned texts, albeit unevenly and not always successfully. These competing entextualizations put forth by scholars are no more disinterested than those they are designed to challenge; they too are sponsored by institutional structures and practices that confer legitimacy and authority. Disputes over entextualizations are therefore not simply

resolved matters of what was said and what was meant. Rather, such conflicts are more fruitfully understood as contests between competing versions of reality, endowed with different degrees and types of institutional authority, and freighted with different interests, goals, and ideologies. Indeed, it is precisely by bringing the private speech of subjects into the public sphere, where it undergoes the entextualizing practices of professional hearing, recording, documenting, and interpreting, that the authority of institutions, whether academic, legal, or political, to make their own reality is established and reproduced.

Appendix: Transcription conventions

.	end of intonation unit; falling intonation
,	end of intonation unit; fall-rise intonation
?	end of intonation unit; rising intonation
!	raised pitch throughout the intonation unit
[]	overlapping speech; numbered brackets mark multiple overlaps in close proximity to one another
< >	researcher comment
—	self-interruption; break in the intonation unit
-	self-interruption; break in the word, sound abruptly cut off
:	length
<u>underline</u>	increased amplitude; emphatic stress; careful articulation of a segment
^	pitch accent
h	exhalation (e.g., sigh); each token marks one pulse
@	laughter; each token marks one pulse
.h	inhalation
#	unintelligible; each token marks one syllable
()	uncertain transcription
word/word	candidate alternative transcriptions

Notes

- * Portions of this article were presented to the Language, Interaction, and Social Organization Research Focus Group at the University of California, Santa Barbara, to the Urban Sociolinguistics Group at New York University, and at the American Anthropological Association meeting in Washington, DC, in the invited session “Transcribing Now: Representations of Discourse in Anthropology,” which I co-organized with John W. Du Bois. My thanks to those audiences for their feedback. I am also grateful to Joseph Park, Peter Tiersma, and two anonymous reviewers for valuable suggestions on earlier versions of this article. Any remaining weaknesses are my own responsibility.

1. It is also worth noting that reality crime television shows that circulate representations of “criminals caught on tape” have socialized viewers into sharing the professional vision of law enforcement officials (Prosis and Johnson 2004; see also Goodwin 1994). Viewers of such shows have become familiar with images of shadowy figures captured in low-quality, grainy video shot from a fixed angle from above (where surveillance cameras are usually located) – and thus learn to think of such images as automatically indicating criminal activity, so that even legitimate action may appear suspicious when seen through the surveillance camera. Such images contrast with the way commentators and other “respectable citizens” are represented on television: they are typically interviewed from an eye-level angle, often looking straight into the camera, with strong lighting, high resolution, and clear sound quality. These representations encourage viewers to trust such interviewees as having a legitimate position in society. I thank Joseph Park for his insights into this issue.
2. I thank an anonymous reviewer for pointing out the legal motivation for this practice.
3. An anonymous reviewer notes that FBI 302s, or summary transcripts of recordings produced using the FBI’s form FD-302, typically no longer make reference to the recordings themselves but only the circumstances under which the evidence was recorded, in order to avoid discrepancies between the 302s and the recordings that could be exposed in court.

References

- Ashmore, Malcolm, Katie MacMillan & Steven D. Brown. 2004. It’s a scream: Professional hearing and tape fetishism. *Journal of Pragmatics* 36. 349–374.
- Bakhtin, Mikhail. 1984. *Problems of Dostoevsky’s poetics*. Minneapolis: University of Minnesota Press.
- Ball, Kirstie & Frank Webster (eds.). 2003. *The intensification of surveillance: Crime, terrorism and warfare in the information age*. London: Pluto Press.
- Bauman, Richard & Charles L. Briggs. 1990. Poetics and performance as critical perspectives on language and social life. *Annual Review of Anthropology* 19. 59–88.
- Blackwell, Susan S. 1996. Corrective measures: Some aspects of transcription in the British legal system. In Hannes Kniffka (ed.), *Recent developments in forensic linguistics*, 255–276. Frankfurt: Peter Lang.
- Blommaert, Jan. 1997. Workshoping: Notes on professional vision in discourse analysis. *Antwerp Papers in Linguistics* 91.
- Blommaert, Jan & Stef Slembrouck. 2000. Data formulation as text and context: The (aesthetics of analysing asylum seekers’ narratives. *Working Papers on Language, Power and Identity* 2. <http://bank.ugent.be/lpi/LPI2.doc> (accessed 3 May 2009).
- Bowe, Heather & Kate Storey. 1995. Linguistic analysis as evidence of speaker identification: Demand and response. In Diana Eades (ed.), *Language in evidence: Issues confronting Aboriginal and multicultural Australia*, 187–200. Sydney: University of New South Wales Press.
- Briggs, Charles L. & Richard Bauman. 1992. Genre, intertextuality, and social power. *Journal of Linguistic Anthropology* 2(2). 131–172.
- Bucholtz, Mary. 1995. Language in evidence: The pragmatics of translation and the judicial process. In Marshall Morris (ed.), *Translation and the law*, 115–129. Amsterdam: John Benjamins.
- Bucholtz, Mary. 2000. The politics of transcription. *Journal of Pragmatics* 32. 1439–1465.

- Bucholtz, Mary. 2001. Reflexivity and critique in discourse analysis. *Critique of Anthropology* 21(1). 157–175.
- Bucholtz, Mary. 2007. Variation in transcription. *Discourse Studies* 9(6). 784–808.
- Cotterill, Janet. 2002. “Just one more time . . .”: Aspects of intertextuality in the trials of O. J. Simpson. In Janet Cotterill (ed.), *Language in the legal process*, 147–161. New York: Palgrave Macmillan.
- Coulthard, Malcolm. 1996. The official version: Audience manipulation in police records of interviews with suspects. In Carmen Rosa Caldas-Coulthard & Malcolm Coulthard (eds.), *Texts and practices: Readings in critical discourse analysis*, 166–178. London: Routledge.
- Cunningham, Clark D. & Bonnie S. McElhinny. 1995. Taking it to the streets: Putting discourse analysis to the service of a public defender’s office. *Clinical Law Review* 2(1). 285–314.
- Daston, Lorraine, & Peter L. Galison. 2007. *Objectivity*. Boston: Zone.
- Eades, Diana. 1996. Verbatim courtroom transcripts and discourse analysis. In Hannes Kniffka (ed.), *Recent developments in forensic linguistics*, 241–254. Frankfurt: Peter Lang.
- Esau, Helmut. 1982. The “smoking gun” tape: Analysis of the information structure in the Nixon tapes. *Text* 2(4). 293–322.
- Foucault, Michel. 1977. *Discipline and punish: The birth of the prison*, Alan Sheridan (trans.). New York: Vintage.
- Fowler, Yvonne. 2003. Taking an interpreted witness statement at the police station: What did the witness actually say? In Louise Brunette, Georges Bastin, Isabelle Hemlin & Heather Clarke (eds.), *The critical link 3: Interpreters in the community*, 195–209. Amsterdam: John Benjamins.
- Fraser, Helen. 2003. Issues in transcription: Factors affecting the reliability of transcripts as evidence in legal cases. *Forensic Linguistics* 10(2). 203–226.
- Gibbons, John. 1996. Distortions of the police interview process revealed by video-tape. *Forensic Linguistics* 3(2). 289–298.
- Goodwin, Charles. 1994. Professional vision. *American Anthropologist* 96(3). 606–633.
- Heritage, John. 1984. A change-of-state token and aspects of its sequential placement. In J. Maxwell Atkinson & John Heritage (eds.), *Structures of social action: Studies in conversation analysis*, 299–345. Cambridge: Cambridge University Press.
- Lyon, David. 1994. *The electronic eye: The rise of surveillance society*. Minneapolis: University of Minnesota Press.
- Lyon, David. 2001. *Surveillance society: Monitoring everyday life*. Buckingham: Open University Press.
- Mann, Steve, Jason Nolan & Barry Wellman. 2003. Sousveillance: Inventing and using wearable computing devices for data collection in surveillance environments. *Surveillance and Society* 1(3). 331–355.
- Mishler, Elliot G. 1991. Representing discourse: The rhetoric of transcription. *Journal of Narrative and Life History* 1(4). 255–280.
- Mondada, Lorenza. 2002. Pratiques de transcription et effets de catégorisation. *Cahiers de Praxématique* 39. 45–75.
- Ochs, Elinor. 1979. Transcription as theory. In Elinor Ochs & Bambi B. Schieffelin (eds.), *Developmental pragmatics*, 43–72. New York: Academic Press.
- Parker, Robert Dale. 1997. Text, lines, and videotape: The ideology of genre and the transcription of traditional Native American oral narrative as poetry. *Arizona Quarterly* 53(3). 141–169.
- Patrick, Peter & Samuel W. Buell. 2000. Competing Creole transcripts on trial. *Essex Research Reports in Linguistics* 32.

- Preston, Dennis R. 1985. The Li'l Abner syndrome: Written representations of speech. *American Speech* 60(4). 328–336.
- Prosisie, Theodore O. & Ann Johnson. 2004. Law enforcement and crime on *Cops* and *World's Wildest Police Videos*: Anecdotal form and the justification of racial profiling. *Western Journal of Communication* 68(1). 72–91.
- Ramirez, J. David, Terence G. Wiley, Gerda de Klerk, Enid Lee & Wayne E. Wright (eds.). 2005. *Ebonics: The urban education debate, 2nd edn*. Clevedon: Multilingual Matters.
- Rice, Tom. 2003. Soundselves: An acoustemology of sound and self in the Edinburgh Royal Infirmary. *Anthropology Today* (19)4. 4–9.
- Shuy, Roger W. 1986. Some linguistic contributions to a criminal court case. In Sue Fisher & Alexandra Dundas Todd (eds.), *Discourse and institutional authority: Medicine, education, and law*, 234–249. Norwood, NJ: Ablex.
- Shuy, Roger W. 1997. Discourse clues to coded language in an impeachment hearing. In Gregory R. Guy et al. (eds.), *Toward a social science of language. Volume 2: Social interaction and discourse structures*, 121–138. Amsterdam: John Benjamins.
- Shuy, Roger W. 2005. *Creating language crimes: How law enforcement uses (and misuses) language*. New York: Oxford University Press.
- Silverstein, Michael & Greg Urban (eds.) 1996. *Natural histories of discourse*. Chicago: University of Chicago Press.
- Solan, Lawrence M. & Peter M. Tiersma. 2005. *Speaking of crime: The language of criminal justice*. Chicago: University of Chicago Press.
- Staples, William G. 1997. *The culture of surveillance: Discipline and social control in the United States*. New York: St. Martin's Press.
- Staples, William G. 2000. *Everyday surveillance: Vigilance and visibility in postmodern life*. Lanham, MD: Rowman & Littlefield.
- Tiersma, Peter M. 1999. *Legal language*. Chicago: University of Chicago Press.
- Walker, Anne Graffam. 1986. The verbatim record: The myth and the reality. In Sue Fisher & Alexandra Dundas Todd (eds.), *Discourse and institutional authority: Medicine, education, and law*, 205–222. Norwood, NJ: Ablex.
- Walker, Anne Graffam. 1990. Language at work in the law: The customs, conventions, and appellate consequences of court reporting. In Judith N. Levi & Anne Graffam Walker (eds.), *Language in the judicial process*, 203–246. New York: Plenum.
- Walsh, Michael. 1995. "Tainted evidence": Literacy and traditional knowledge in an Aboriginal land claim. In Diana Eades (ed.), *Language in evidence: Issues confronting Aboriginal and multicultural Australia*, 97–124. Sydney: University of New South Wales Press.
- Walsh, Michael M. 1999. Interpreting for the transcript: Problems in recording Aboriginal land claim proceedings in northern Australia. *Forensic Linguistics* 6(1). 161–195.

Mary Bucholtz is a professor in the Department of Linguistics at the University of California, Santa Barbara, where she specializes in sociocultural linguistics with a particular focus on language, identity, and power. Her research on institutional and cultural practices of linguistic representation, including transcription, appears in *Discourse Studies*, *Journal of Pragmatics*, and *Journal of Sociolinguistics*, among others. Address for correspondence: Department of Linguistics, 3607 South Hall, University of California, Santa Barbara, CA 93106-3100, USA <bucholtz@linguistics.ucsb.edu>.