Applied Linguistics 2022: 1–22
https://doi.org/10.1093/applin/amac005

For the Record: Questioning Transcription Processes in Legal Contexts

*Emma Richardson, Kate Haworth, and Felicity Deamer

Aston Institute for Forensic Linguistics, Aston University, UK
* Email: e.richardson4@aston.ac.uk

Written records of spoken interaction are typically assumed to be adequate for the purpose they serve, often receiving minimal scrutiny from the institutions which consume them. In this article, we scrutinize the current practices of capturing spoken interaction in legal contexts in England and Wales, and highlight some of the often serious legal consequences that result. We ask five questions of record keeping in legal settings: (i) Is the record produced an accurate representation of the spoken interaction?; (ii) Do lay and PPs have ownership? Answered by giving careful thought to the rights they may or may not have to their data; (iii) Who has agency, who's 'voice' is represented in the recorded account?; (iv) Then, we ask how usable the record is; and (v) How resource efficient it is to produce and use. By asking these questions, we make visible the underlying assumptions about transcription adequacy—in doing so, we acknowledge and enable reflection on the process of capturing spoken interaction. We envisage this model to be applicable to a range of institutional settings.

INTRODUCTION

Written records of spoken interaction play a central role in many professional contexts. Across a range of institutional settings, 'practitioners', or 'professional participants' (PP), elicit and capture spoken talk from various 'clients' (Sarangi 1998) or 'lay participants' (LP), create records of that talk, and later refer to the record in place of the original interaction (see Haworth 2018). This process of transcription, a form of entextualization (Bauman and Briggs 1990; Park and Bucholtz 2009; Maybin, 2017) removes language from one context to be used by another party at a later date, in another setting.

Writing in the late 1970s, Ochs focussed on the selective process of transcription, exploring how the product, the resulting text, reflects a set of theoretical goals and definitions (Ochs 1979). At the time, there was relatively little empirical focus on transcription of the spoken word other than in the emerging discipline of conversation analysis (Sacks et al. 1974). Since then, many academics have studied the transcription process in a range of contexts (Jefferson 1985, 1996, 2004; Biber 1988; Mondada 2007) including a consideration of how we transcribe multimodal text-based data (Meredith 2016: 253). However, these developments have not (yet) translated over to

improvements in transcription practices in institutional contexts. Fifteen years ago Blackwell (1996) stated that 'the need for forensic linguists to develop an understanding of the transcription process is as pressing as ever'. Yet, in the early 2020s there is little recognition in legal contexts of the long realized basic principle of the non-equivalence of spoken and written text (Biber 1988; Halliday 1989), and hence minimal scrutiny of the process of converting one to the other (see e.g. Haworth 2018). We are currently conducting a wider collaborative project aimed at bridging that divide between academic expertise and professional practice, and in the course of this we have undertaken considerable reflection on what shapes institutional norms in converting spoken data to a written record. The outcome of those reflections is presented here.

It is important to acknowledge at the outset that written versions of spoken data are necessary. It is equally important to start from the premise that no transcript of spoken interaction can be exact. Transcripts are only ever a representation of the spoken talk and never direct copies, and they inevitably result in a loss of detail. However, a key issue is the extent to which even this basic premise is not recognized within the institutions which produce and use these records. A written record (however detailed) can never capture spoken interaction in its entirety, but it can, when purpose and function are carefully considered, be fit for purpose.

We set out a series of questions to be asked of any transcript of spoken data, aimed at assessing whether it is fit for its intended purpose. In order to frame and illustrate the discussion, we apply our proposed approach to a specific type of data, namely official records of police interviews in England & Wales (E&W). This data type is of particular interest, not only because of the very serious uses to which such records are put, but also because of the many different formats in which the data are produced through the criminal justice process. The fact that records of data as similar as police—suspect and police—witness interviews are produced so differently, provides us with an ideal context in which to demonstrate the factors at work, and to test our theories as to how best to approach issues of adequacy and accuracy. However, we envisage that our model can be applied equally well to written transcripts of spoken data across any context, including our own professional context as linguists who analyse spoken text. It thereby also illustrates how, by attempting to address a 'real-world' practical problem, we can in fact learn a great deal about our own methods and academic practices.

WHY PRODUCE WRITTEN RECORDS OF SPOKEN INTERACTION?

Spoken utterances, once issued, are lost; if their content needs to be preserved for any reason, some form of record must be created. There are many institutional and professional contexts in which this therefore becomes necessary,

such as in medical contexts where records of interactions with patients may form a key point of reference in decision-making processes; in the UK parliament where the interaction that takes place is formally recorded in Hansard (see Slembrouck 1992; Mollin 2007; Cribb and Rochford 2017) or when asylum seekers are interviewed by officials in their non-native language as part of the application process (Blommaert 2001). As academics, researchers, and certainly as linguists, we also produce transcriptions of spoken data and later use them as evidence to support analytic or theoretical claims (Bucholtz 2000, 2007, 2009). It is important to note that the stakes for the participants are likely considerably lower in some settings, such as research interviews than others, such as suspect interviews in murder investigations.

Spoken utterances in an institutional setting, essentially when a LP is producing a narrative (relaying an event or a series of events) to a PP, are generally recorded in order to serve an institutional purpose. By PP we mean a person within the interaction who is participating as part of their professional capacity, and therefore usually with expert knowledge of the institutional context within which the interaction is taking place. By LP, we mean those who are not part of the institution, and whose participation in this particular interaction is not in a professional capacity. Some examples from the legal context would be a witness, victim, or suspect of a crime (as LP) relaying a version of events to a police officer (as PP) in a formal investigative interview or, a defendant (as LP) being questioned by a barrister (as PP) in the courtroom. Records of these interactions are routinely produced, in order to render them accessible to future audiences within the legal process (Walker 1990); these records are criminal evidence in some cases.

One of the central functions of the record in legal settings involves various PPs investigating and testing the recorded account, including comparing it both with accounts obtained from others, and with accounts given by the LP at a different point (to a different audience, or in a different context, for example). At the account elicitation stage, the PP often has a far greater knowledge of the full extent of who will make use of the records of what was said than the LP (Haworth 2013). Maybin (2017) states that 'knowledge about prior and current contexts is only ever partially shared in professional encounters and recontextualization is not equally transparent to all participants' (p.421). The 'textual trajectory' (Blommaert 2005) the record takes often foregrounds the voice of one party (PP) and marginalizes another (LP). The 'institutional power' PPs have in the process (see e.g. Bauman and Briggs 1990; Linell 1998; Blommaert 2005; Park and Bucholtz 2009) is worthy of further exploration, especially when the transcripts produced serve a legal evidential purpose, and we seek to shed further light on this here.

POLICE INTERVIEW DATA TYPES

The scope of this discussion is deliberately limited to only one type of official record of spoken interaction, namely interview records, and predominantly

one jurisdiction (E&W), although comparisons are made with other jurisdictions where appropriate. An essential distinction to draw here is between police interviews with suspects and victims. For each, a police officer (PP) is instigating a dialogue with a person who has a connection with an alleged crime being investigated with the purpose of gathering information and (potentially) evidence for that investigation. However, there are many important differences between interviews with suspects and with victims, most notably how the interaction is captured in the official record and how those records are later used. We will therefore consider each in turn.

SUSPECT INTERVIEWS

The first data-type for a suspect interview is the audio or video recording. Investigative interviews with suspects are routinely audio-recorded (Police and Criminal Evidence Act 1984, Code E) with only a few exceptions (e.g. some terrorism cases). Until relatively recently, the medium for recording police interviews in England and Wales was a cassette tape, although they are now largely digitally recorded. Increasingly, suspect interviews are also video-recorded, although this is largely reserved for the most serious cases, and is generally a matter of discretion. Cassette tapes formed a physical piece of evidence, present in police and prosecution case files; digital recordings are a less tangible artefact.

The second data type is the written version of the audio (or video) recording, for which there are several options available. For the most serious cases, a 'full' transcript may be produced, which aims to encapsulate the entire interaction, including every speaker turn. Far more common is the production of a much shorter, abridged version of the interaction, generally known as a ROTI/ROVI, or 'Record of Taped/Videoed Interview'. These aim to record the most important parts of the interview 'verbatim', with other parts summarized or glossed over. A final option is to produce only a short summary interview record. This might happen, for example, where there is little to no likelihood of the matter ending up in court, and so no official record is likely to be required. Since these generally remain in an internal police or Crown Prosecution Service file, they have been subject to little academic scrutiny.

During the investigative (pretrial) stage, interview data are therefore available in two formats: the audio/video recording, and the written transcript of that recording. If the interviewee is then formally charged with a criminal offence, and the case goes to court, the police interview will be formally presented to the court as part of the prosecution case against them. Typical practice is for the interview record, usually the ROTI, to be read out to the court. This produces a third data-type as the record is converted back into an oral format. The written format is used rather like a script, performed by

members of the prosecution team, usually the prosecution lawyer and a police representative (see Haworth 2018).

WITNESS INTERVIEWS

For witnesses, the process is rather different. Standard practice is to produce a 'witness statement', rather than an interview record. Witness interviews are not routinely audio or video recorded; instead, the interviewing officer produces a written, monologic summary of what the witness says, from the responses given to the interviewer's questions and other prompts. The summary is constructed during the interview itself, and negotiating what goes into this record forms a key part of the interaction (see Rock 2001). The resulting witness statement is used for the investigation, as with a suspect interview record, but it has a different status when it comes to court. Unlike suspect interview records, witness statements are not routinely introduced as evidence. Instead, the witness is called to give their evidence in person (their 'evidence-in-chief' in court). However, if the in-person account differs in any material way from what is in their witness statement, then the opposing lawyer can apply for the witness statement to be introduced in evidence as a 'prior inconsistent statement' (ss. 119 and 124 Criminal Justice Act 2003); at this point, the witness statement does become an evidential document. It is certainly not uncommon for a witness to be cross-examined quite thoroughly about the contents of their witness statement.

An important exception to this standard procedure is for witnesses who are institutionally defined as vulnerable or intimidated (ss. 16 and 17 of the Youth Justice and Criminal Evidence Act 1999). The 'special measures' available to such witnesses include allowing a 'video recording of an interview of the witness' to be admitted as their evidence-in-chief, removing the requirement for the witness to give evidence in person in court (s. 27 YJCEA 1999, with further developments currently in process). This is clearly a substantial departure from the processes described above, and hence makes an interesting point of comparison.

FACTORS INFLUENCING THE FORMAT OF EVIDENTIAL RECORDS

We can see, then, that interview records are produced in different formats and with different methods, which leads us to the question of *why*. It seems to relate partly to the type of interaction, but more importantly to the later uses to which the record will be put. The evidential use of the record further down the line could be argued to directly influence the choice of recording format and transcript style. However, on closer examination, it becomes clear that it is not so simple; nor does that neat correlation necessarily hold up. Witness interview records

may not be intended to be routinely produced in court like a suspect interview record, but in reality they often do end up being introduced as evidence.

There are also several historical and traditional reasons seemingly behind some of the differences. For example, audio-recording of suspect interviews was introduced by Police and Criminal Evidence Act 1984 (PACE) in order to address serious problems of misconduct, forgery, and even violence against suspects during interview (Bucke et al. 2000); witnesses were presumably not seen as requiring such protection and so the recording requirement was not extended to cover witness interviews. In England and Wales courts, there is a general principle of orality, so all evidence is routinely introduced orally before the court; this may explain the practice of reading suspect interview transcripts out loud. It also must be said that change often happens extremely slowly in the legal process; it is entirely plausible that suspect interview records are read out loud (rather than, for example, playing the recording) simply because they always have been, and nobody has seen reason to change.

We consider it well overdue to take a fresh look at interview records, and in the sections that follow we make an attempt to set out the factors which we consider ought to be taken into account. We start from the position that the most meaningful question to be asked of a transcript is: is it fit for purpose? That purpose may vary considerably, and that is indeed why different types of transcripts may be perfectly adequate, depending on what they will be used for (see e.g. Cook 1990; Lapadat 2000). We therefore consider it essential that transcripts should be measured against that specific purpose. In order to do this, we develop certain principles when producing records (see Meredith 2016) and propose the following series of questions, to be asked of any transcript, and answered with reference to that purpose:

- How close is it to the original?
- Who has agency?
- Who has ownership?
- How useable is it?
- How much resource does it require?

This is not intended as a definitive list; instead we encourage the reader to consider them as overlapping, intersecting, and developing areas worthy of further investigation. Our intention is to make visible some of the issues, and to invite readers to apply them beyond the England and Wales jurisdiction and beyond the legal setting. Later, we consider how record types emerge at the intersection of these concepts, by applying this framework to specific types of interview record.

The data cited below are from two sources: the data set collated by Haworth for a previous study (Haworth 2009); and the official transcripts of the trial of Harold Shipman. Shipman was a British medical doctor, who was convicted in

2000 of murdering 15 of his patients; a subsequent public inquiry concluded that the total number of victims was nearly 250. The publicly released inquiry documentation includes the full transcripts of his trial.[1]

QUESTION 1: HOW CLOSE IS IT TO THE ORIGINAL?

The first question we propose to ask is how close the data type is to the original spoken interaction. In conceptualizing this criterion, we consider it important to distinguish it from the question of accuracy: inaccuracy implies mistakes or errors, whereas what we are referring to here is more a question of how well the transcript captures the entirety of the original data. An audio or video recording is closer to the original interaction than a written transcript; it inevitably preserves features such as accent, speaker emphasis, paralinguistic features, embodied actions, and so on (more on which below). However, this is only the 'broad-brush' picture; the realities are often rather more nuanced.

When we consider how close police interview records are to the original data, several studies have revealed serious problems in this respect. As previously mentioned, audio recording of suspect interviews was introduced to counteract historic problems with the production of highly inaccurate interview records (Dixon and Travis 2007); including cases where interviewers writing up the official record were found to have entirely fabricated a suspect's confession (see e.g. Coulthard 1996, 2002). Aside from such extreme examples of deliberate abuse of the written record, the introduction of audio recording also addressed the highly problematic practice of the officer being expected to write down what happened post interview from memory (see e.g. Coulthard 1996, 2002), a practice that inevitably resulted in the loss of substantial amounts of the original interaction. In a study of similar practices in Sweden, Jönsson and Linell (1991) found that police-written reports were shorter, denser in information, included more complex noun phrases, and were more structured than the oral telling (see also Komter 2019).

For witness statements, produced by the interviewer during the interaction, the dialogic, multi-participant interaction is transformed into a monologue, thereby completely erasing the input of the interviewer and rendering it inaccessible to all future audiences. Rock (2001) demonstrates the transformations which consequently occur in the witness's version of events, from their original telling, through various iterations as the interviewer probes and shapes their account, to the final version recorded in the official statement.

The audio recording of UK police–suspect interviews is clearly a considerable improvement on previous practices. However, it is important to sound a note of caution regarding the extent to which even an audio recording is an accurate encapsulation of the original context. Ashmore et al. (2004: 361) highlights the dangers of treating audio recording as 'a direct and evidential record of a past event, and thus as a quasi-magical time machine' drawing attention to the mediating role of the recording between the hearer and the original context.

Audio/video recording has also not replaced the need for a written transcript to also be produced (see further below), and in practice, it is the transcript which is relied on rather than the recording (see Haworth 2018). Yet the inevitable distortion which occurs in the process of transcribing these recordings into an official transcript continues to be overlooked. The types of feature which are lost when audio or video data are converted to a written transcript (e.g. pauses, gesture, pace, and intonation) have been well described elsewhere, so we do not propose to set these out again here (see e.g. Walker 1990; Shuy 1993; Gibbons 1995, 2001; Eades 1996; Fraser 2003, 2014; Bucholtz 2007, 2009; Komter 2012).

In the police interview context, the written record's future role as evidence in fact results in some para- or extra-linguistic features (Walker 1990) actually being better preserved than with many audio-only recordings, since any embodied conduct that is deemed relevant by the interviewing officers may be reformulated as spoken talk as 'for the benefit of the tape' (Stokoe 2009), in other words for the future audiences who will consume the interview data (Haworth 2013). However, far more features will not be captured in this way, and are therefore lost before the transcription process takes place.

In terms of selecting which features are preserved, processes and practices surrounding transcription in almost all contexts are typically to do with focus; those elements of speech which are considered 'relevant', by the transcriber within a certain setting, are transcribed. For police interviews there is a focus on content, to ensure all or most of *what* was said is transcribed, with much less emphasis on *how* it is said. This is in contrast to how a linguist might produce a phonetic transcription of the speech for example. Choices of 'relevance' are partly due to practical considerations of readability and usability (see further below); but it also appears to be symptomatic of a lack of recognition within the legal system of how much meaning is conveyed by these features.

QUESTION 2: WHO HAS AGENCY OVER WHAT GOES INTO THE RECORD?

There are two distinct aspects to consider when addressing this question. For the first, we use the notion of agency to refer to the extent to which the LP's 'voice' is represented in the telling of the account. For both suspect and witnesses, their account of the events is tightly controlled through the question and answer interview format. The literature on interview interaction has demonstrated how interviewing officers (PP) and the interviewee (LP) collaboratively produce the LP's account, but with the result that the LP's account is heavily restricted and influenced by the PP's agenda (e.g. Auburn et al. 1995; Heydon 2005; Haworth 2017; MacLeod 2020). This literature highlights how

the process of elicitation impacts negatively on the agency the interviewee has over what *can be* said and ultimately therefore what *can be* recorded.

While acknowledging the importance of that point, for the purposes of this article we wish to extend this topic by adding a focus on agency over the official record. What emerges very clearly from such a focus is that generally speaking, once LPs have spoken in the interview room they lose virtually all agency over what happens to their talk. We seek to question whether this is the most appropriate balance; or rather, we seek to encourage those with control over this factor to ask this question.

This factor relates closely to the following topic of ownership; in our treatment of these factors, we consider agency in terms of what goes into the official record, then move on to the question of ownership over those records once produced.

Suspect interviews

Starting with police–suspect interviews, the first point to observe is that the act of recording is entirely controlled by the police PP in the room, including when to start and stop the recording. These decisions frame what 'counts' as the interaction, and of course what is available to be transcribed. It is worth remembering that there will of course be talk before and after these points (see Komter 2019 and Arnold 2021 for a researcher's rare observation of this, and the significant consequences for the interview itself), however the police interviewer gets to act as sole gatekeeper of when the official, institutionally recognized interaction begins and ends. (We refer here only to the start and end of the recording; pausing recording *during* the interaction would be a breach of PACE 1984 and would likely render the entire interview as inadmissible.) This can be observed in the following example:

Example 1: Interviewer closing the interaction.

```
1     IR:   a decision will be made. okay! (.) we'll bring the
2           interview to a close,
3     IE:   and I want a solicitor [ cos ] I'm not even taking this
4     IR:                        [yeah.]
5     IE:   shit.
6     IR:   no problem. (.) okay [the time]
7     IE:                      [ seven  ] witnesses again[st one. ]
8     IR:                                                [the time]
9           (.) the time is now, (-) one (-) oh four. (.) p m, (-)
10          this is a form G (.) fifty three, (-) which basically
11          tells you what's happening to your tapes, (.) so I'll give
12          you that. (.) okay? (.) again before we close, (.) have
13          you got any (.) complai[  nts  ] about- okay. (.) lovely.
14    IE:                        [no(ne.)]
15          (--) one oh four, (.) p m, (.) interview's finished.
```

(Haworth, IV 5.11.2/1: 609–20)

Further, the official transcripts, in whatever format, are produced on behalf of the police, by police employees (see Haworth 2018). In other words, prosecution PPs have sole agency over what is preserved in the official record, including the editing process involved in the production of ROTIs. The LP and their legal team (henceforth defence PPs) are not consulted, and do not have the right to be involved in the production of the official record.

Witness interviews

For witness interviews, agency similarly rests solely with prosecution PPs, although with important differences. Here, the written record is produced as the interaction takes place, with the PP constructing the record while the witness is present. Although this would appear at face value to be a collaborative process, with the LP first-person monologue format implying a high degree of LP agency, in practice the PP has full control over what goes into the record. At the end of the process, the witness will be asked to sign this statement, confirming 'I believe that the facts stated in this witness statement are true'. This would appear to give some agency to the LP in terms of being able to correct or amend the official version, but in practice this also seems to be a rarity, presumably at least in part due to the power dynamics at play (making it socially and interactionally difficult for a witness to challenge and correct their police interviewer), as well as the likelihood that the LP (i) does not recognize the significance of what exactly is in the record, and how it might subsequently be used (witnesses are not cautioned); and (ii) an erroneous assumption that the statement will be an accurate record of what they just said, underestimating the immense difficulty of this task. This confirmatory signature, on a statement written with LP as first person, is thus potentially more dangerous, in that it gives the illusion of agency for the LP, when the reality is that they had little, if any. The process leaves scope for institutions to encode alternative, congruent, narratives in the statements, as Canning (2018, 2020) has shown in an analysis of statements given in response to the Hillsborough Disaster, where 96 Liverpool fans lost their lives at a football match due to operational failures. The consequences for witnesses of the lack of agency can be seen further down the line in the criminal justice process. As Rock (2001: 47) notes, 'the statement is not used simply to support, confirm or even explore the witness' presentation of events, rather it becomes an authoritative text against which even the witness, the original source of the information contained in the statement,

is assessed'. This can be seen in action in the following example of a witness being cross-examined at trial:

Example 2: Cross-examination of John Alan Green at the trial of Harold Shipman, Thursday, 14 October 1999 (official court transcript).

DC: Is that the statement you made to the police on the 4th August 1998?
W: Apart from the end bit.
DC: I am sorry?
W: Apart from the end bit. [...]
DC: Are you saying that is not what you said to a police officer on the 4th August of 1998?
W: Yes. […]
DC: Did you sign that statement?
W: Unfortunately yes. [...]
DC: If you didn't say what you now maintain was correct, why did you sign the statement?
W: I was probably worked up, I don't know. […]
DC: When you made that statement was the police officer asking you questions, you would give an answer and the police officer would write it down, is that how it was done?
W: Yes.
DC: Were you given an opportunity to read over the statement before you signed it?
W: Yes.
DC: Did you read over the statement before you signed it?
W: I did your Honour, yes. […]
DC: That was the statement that you signed on the 4th August as being true to the best of your knowledge and belief?
W: Yes.
DC: Are you now saying, Mr. Green, that that sentence is in fact incorrect?
W: Yes.
DC: Well, if you are now saying it is incorrect, why did you allow that statement to go through signed by you?
W: Couldn't he make a mistake, the gentleman writing it down? I mean we all make mistakes. [...]
DC: Are you now saying, Mr. Green, that the police officer made a mistake in writing down your answer?
W: Yes. [...]
DC: And you signed it?
W: I did.

At the court stage, witnesses (including defendants) regain some agency over their account, indeed this is the first time since the original interaction that the LP is provided an 'on-record' opportunity to comment on what is being represented as their account. Yet this is still mediated and constrained by the questioning of PPs, in the form of lawyers for both prosecution and defence, who thereby largely retain control over the witness (see e.g. Cotterill 2003). Further, the ability of a witness to reframe, or exercise any form of control over their version, is still heavily curtailed by what is written in the official record. As this example demonstrates, they are measured against it, and any discrepancies attributed to a failing on the part of the witness, not the record. This is therefore a very limited form of agency, whereby the act of reclaiming their account in their own words can actually do them substantial harm. This is highlighted in a further exchange between this witness and cross-examining counsel:

Example 3: Cross-examination of John Alan Green at the trial of Harold Shipman, Thursday, 14 October 1999 (official court transcript).

```
1          [Defence Counsel reads part of his statement]
2    DC:   What is the next part of the sentence please?
3    W:    "That she had heart attack."
4    DC:   "And informed us that she had had a heart attack," is that
5          correct? Is that what is in your statement?
6    W:    Yes. [...]
7    DC:   You use the words "heart attack" or those are the words
8          contained in your statement?
9    W:    Yes.
10   DC:   Today you were adamant that the words used to you were
11         "cardiac arrest?"
12   W:    Yes.
13   DC:   Which is correct, Mr. Green?
14   W:    Cardiac arrest. That's definite. I don't want anybody
15         twisting my arm.
16   DC:   So was this another example of the police officer not
17         correctly writing down what you had said?
18   W:    Yes.
```

We must of course acknowledge that the version presented here is mediated through the official court transcript, however the damage that counsel is attempting to inflict on the witness's credibility is clear. It is also not stretching the bounds of interpretation to identify the note of institutional incredulity that a police officer might not accurately record what a witness said, reflecting the faith in the official record which seems to be prevalent in the current system. What is also interesting, is that this same counsel recognizes the power of a witness presenting their evidence themselves, by having them read out the specific sentence of the statement to which they wish to draw attention, exploiting agency and ownership.

A practice has developed which does tip the balance of agency towards the LP, however. This entails a suspect giving a prepared statement at their police interview, usually accompanied by a refusal to engage with the interviewer's questions. This is a fascinating shift in agency from prosecution (police) PP to LP. Firstly, it completely sidesteps the constraints of the institutionally controlled Q&A, enabling the LP to give an unfiltered, direct account, while still maintaining cooperation with the interview process. Secondly, it enables this LP-controlled account to be recorded as the official version, in that its delivery will be audio-recorded, and, presumably, transcribed into the official transcript (although this requires further research for confirmation). That defence lawyers have developed this practice tells us a great deal about the importance of agency over LP's accounts in this context. However, it must be acknowledged that such a tactic is in practice largely the preserve of those with sufficient personal funds to access private legal representation (linking with the question of resource discussed below).

QUESTION 3: WHO HAS OWNERSHIP?

We now move on to consider a related but separate point, regarding ownership of the interview record. In the police interview context, the LP has no ownership whatsoever over the official interview record; all versions of the data are owned and controlled exclusively by the PP. Audio or video recordings, and their accompanying official transcripts, are classed as prosecution evidence. Indeed it is this ownership of the record that allows only the police/prosecution PP to produce the official transcript, and hence to select what goes into it.

Copies of these items of prosecution evidence are provided to the defence through the usual process of disclosure. If an LP or their legal representative wishes to challenge or amend the contents of the official record, this then becomes a complex matter of legally challenging the prosecution evidence.[2] In practice, however, this rarely seems to happen, or at least not on the grounds of accuracy or representativeness.[3] The defence could also have their own transcript produced from the audio recording, if requested, although it would not be an official version, and again they would need the permission of the court to introduce this version in evidence. But as with many other aspects of this process, the cost is likely to be prohibitive, especially for LPs reliant on legal aid. (At this point we should note the role that many linguists have played in giving expert evidence on disputed audio and transcripts, however to our knowledge that has rarely, if ever, happened with a police interview record and is beyond our scope here.)

The Prosecution PP ownership of the record is also the reason why, at the court stage, the transcript is read out loud solely by Prosecution PPs (Prosecution Counsel and a police representative). This is highly problematic, in that this conversion back from written into spoken format inevitably involves further transformation of the data, yet from an entirely prosecution-driven perspective with all the potential bias that entails (see Haworth 2020: 148–9 for an example). However, this therefore gives rise to a relatively easy 'fix'; agency could be returned to the Defence at this stage, by having a member of the Defence, potentially even the LP (interviewee) themselves, reading out the LP's turns in the interview record.

In considering ownership of records of spoken interaction, the police interview context likely represents an extreme; it is also easy to understand why a PP ought to retain ownership in terms of overall control of criminal evidence. But we suggest that this is why *agency* should perhaps be compensatorily tipped rather more towards the LP, especially in terms of the content of the official record. Those with ownership of such data, including ourselves as academic researchers, might consider the impact of that ownership in our relationships with our own LPs, and how ownership and agency should be balanced and distributed. By framing this as a question to be asked of our data records, no assumptions are made as to how these factors should be weighted; this will vary according to circumstance. Our approach seeks to allow for the

fact that there is not only one acceptable answer, while nevertheless emphasizing that these questions still need to be asked, and active decisions made as to where the balance should lie. However, while making those decisions, it is also essential to factor in the practicalities of the context, and that is what our final two questions address.

QUESTION 4: HOW USABLE IS IT?

Any consideration of this topic must take into account the highly practical factor of how easily the record can be used for its intended purpose. Using an audio or video recording might be considered ideal from the perspective of keeping the data closer to the original, however, at the time of writing the audio/video record as a data type is less usable than the written counterpart in the legal context. By usable we essentially mean two things: how 'portable' and 'referenceable' a version is. To unpack this further, interview records are used by a wide range of professional recipients in the legal system, especially during the police investigation and any subsequent court hearing (see Haworth 2013). The record needs to be easily shareable, and easily consumed by those it is shared with. A cassette tape, disc, or digital file requires equipment to be able to access the data, while a paper copy can be viewed immediately and without equipment, making it more portable in *most* legal contexts. It is therefore easy to see why written transcripts are relied on so heavily in case files, court hearings and investigations.

The second term, referenceable, refers to the ways in which users are able to locate, mark and reference part of the material from the whole. It also must be acknowledged that while a user might note down where within an audio file they would like to refer to, either by noting start and end times, or making a shorter clip, there is no doubt that highlighting a text or noting a page number for all who have access to the text is easier. This can be seen in the following examples, taken from *R v Shipman*, where specific exchanges during interview are referred to in court by their page number and position in the written text:

Example 4: Cross-examination of Shipman, trial day 33.

```
Q:   You see, I am going to suggest you were never ever justified
     with the state of your knowledge of Kathleen Grundy in ever
     putting old age on that death certificate, never. I am going
     to ask you please to look at your interview, interview bundle
     please members of the jury, what you said about old age to
     the police, page 44 of the bundle the last two lines of page
     44. "What opinion would you give with hindsight about Mrs.
     Grundy's general state of health? …"
```

This is, of course, always under review, and while technological developments that would tip this balance exist, which would make portability and referenceability just as easy if not more so in the digital audio format, these are yet to be implemented into practice in this setting.

Another aspect here is that those working in the legal system are always heavily pushed for time. A written record of any kind is quicker to scan through than having to listen to a full interview; similarly they generally do not have time to read through lengthy interview records to glean the most relevant and important points contained in them, so a 'full' transcript is often actually less usable in practice than a summary/ROTI. When it comes to presenting the interview as evidence in court, however, the examples of transcripts being (mis-)read out loud discussed in Haworth (2018: 443–45) demonstrate that this method can actually be *more* time-consuming than simply playing the recording, raising further doubts over this practice.

Finally for this question, assuming for now that producing a written record is necessary, then we must consider what level of transcription detail can be considered as usable by the actual audiences. Gibbons (2003: 30) states 'If a transcript cannot be understood as readily as the oral language that it represents, then it is failing in its primary task of communicating what was communicated in the primary context'. The more detail of the spoken interaction represented through transcription conventions, arguably the less usable the transcript becomes. Linking back to our first question, this is where there is a trade-off between accuracy and usability. As linguists producing transcripts, we might consistently include overlapping talk, pauses, lack of fluency, stress, intonation, and so on, in order to better encapsulate the original interaction, on the basis that without denoting these features, critical components of the interaction can be lost (Walker 1990; Jefferson 2004). However, the level of detail that is included in a transcript for discourse analysis comes at the expense of the readability for a general audience (Fraser 2014), and transcripts that are used in legal contexts are not required to be of the standard required for a linguistic analysis (Blackwell 1996). Lengthy, detailed transcripts place a burden on readers, demanding extra time and energy from them; interpreting some of the features included in a linguistic transcript may require expertise beyond the scope of typical participants in the legal system. This is therefore once again a question of balance, and, crucially, matching the transcript to the intended purpose (Cook 1990; Lapadat 2000).

QUESTION 5: HOW MUCH RESOURCE DOES IT REQUIRE?

Thus far, we have to some extent worked on the assumption of an ideal situation, where records *can* be produced to the standard required for the purpose. However, when considering 'real-world' applications such as this, it is essential to take into account the practical realities of the context, and by far the biggest factor in this context is that of resource. It is of little use producing

recommendations if it simply not possible to implement them; it is equally unfair to level criticism at those doing their best while working within considerable restraints. It therefore has to be acknowledged that what can be achieved, in terms of producing transcripts which are fit for purpose, ultimately involves a trade-off between what is ideal and what is practical. In the police interview context on which we have been focusing, the main resource impact is in (i) staff and (ii) time, both of which are in short supply. In order to produce the volume of interview transcripts required daily, police forces generally employ a pool of 'ROTI typists' (their exact title varies from force to force), whose sole job is to convert audio or video recordings of interviews into official police transcripts. There appears to be generally very little financial resource available for this; the number of such staff also appears to vary widely from force to force, presumably in accordance with both budgets and priorities (we are currently conducting research which aims to shed much more light on this less acknowledged work, although see Haworth 2018).

In terms of time, police transcripts generally need to be produced in a very short time-span, certainly much shorter than any linguist would be comfortable taking over our own data transcripts. Transcripts of interviews will need to be available for reference by the ongoing investigation, hence are often required very swiftly after the interview has taken place. Due to these two very real factors, decisions will have to be made as to what goes into the transcripts, weighed up alongside all the other factors we have discussed above. It is immediately obvious that 'full' transcripts cannot be produced for all interviews; there simply are not the resources for this. But equally, a full transcript may well not be required in the context; for example it may already be agreed that no further action is going to be taken, or it may be a very minor offence with no facts in dispute, in which case it would be a waste of precious resource.

Another factor to have in mind is that, once we reach the court stage, it is no exaggeration to say that time is money; court time is extremely expensive and so it is important that time is not wasted presenting more detail than is necessary for the case in hand (it can be seen that this links strongly with the 'useability' question above). At all stages of the criminal justice process, then, resource efficiency is a major factor in the level of detail and time put into transcripts. However, questions remain as to who makes that judgement, and at what cost to justice?

FINDING THE RIGHT BALANCE

We have set out five concepts, identified from the literature and from our own research into transcription in legal contexts, which we consider to be the most important deciding factors in determining the type of written record of spoken interaction which will be most appropriate for the given purpose. These are (i) accuracy, (ii) ownership, (iii) agency, (iv) usability, and (v) resource efficiency. For any given transcript, each of the concepts has the potential to score

high or low. Consider the scale as the authors marking them on a Likert scale where the number 10 represents total or complete accuracy, ownership, agency, usability, and resource efficiency. How they score is often a trade-off, since when one is scoring high another might be scoring low as a result. We envisage that these five concepts can be mapped onto a radar diagram (Figures 1-3 below), as a means of visually representing the balance that needs to be struck, as well as identifying instances where they are out of balance. To illustrate this, we will now apply them to several police interview record types in turn.

Audio/video recording of a police-suspect interview

Audio/video recordings score high for accuracy, in that they preserve the original interaction, albeit in a somewhat impoverished form (Figure 1). They score mid-range for LP agency, in that the LP does not have agency over the production of the recording, but does at least know that the recording is taking place and has some opportunity to say what they wish on that record. Ownership scores low from the LP perspective, as with nearly all legal evidential records. Finally, the audio/video format scores low for usability, and for resource efficiency (especially in terms of how they can be presented to later audiences). This illustrates that the increased accuracy comes at a price in this context.

Witness statement

Witness statements score much lower for accuracy (Figure 2). They also score lower for agency, in that they have much less opportunity to have direct input into what is recorded in the official version. They share the same low ownership as with other police interview records. However, they score highly for usability. They also score highly for resource efficiency, in that they require no

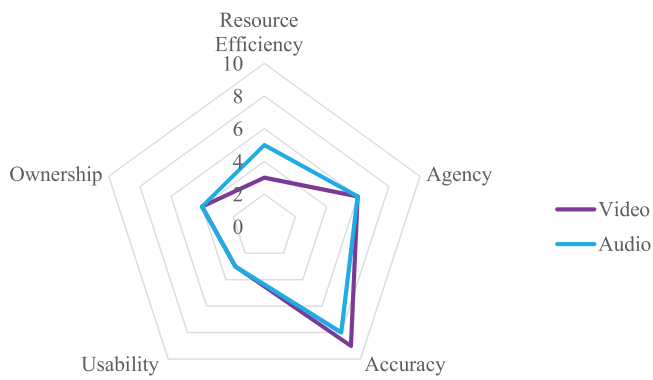

Figure 1: Video and audio recordings.

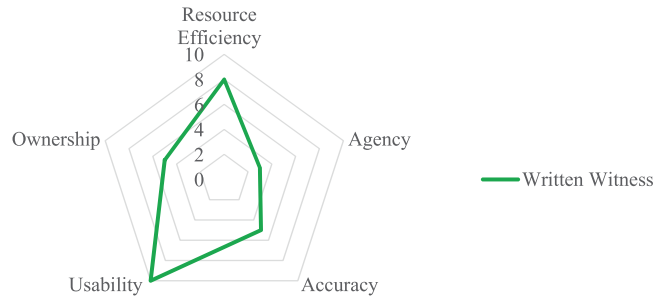

Figure 2: Written records.

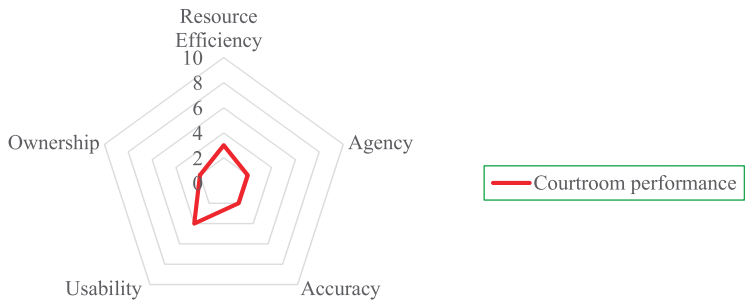

Figure 3: Oral courtroom performance.

hardware, and only take up the resource of the interviewer's time in conducting the interview and simultaneously writing the formal record.

Courtroom 'performance' of suspect interview records

Figure 3 depicts a particularly telling example. The performance of a record scores low for accuracy, and low for LP agency and ownership. It must also score low for usability; indeed this performance is less 'useable' than the transcript which it is based on. In terms of resource efficiency, it is possible that it is less resource-intensive than playing an audio/video recording, however it surely takes up just as much court time, which is arguably the larger spend, meaning that in reality this is comparable to a recording. This analysis therefore indicates that there is little to recommend this practice.

DISCUSSION

Written records of spoken interaction in legal contexts have seemingly resulted from a more or less organic evolutionary process, where considerations of current and/or future purpose, together with various assumptions and influencing factors, have worked to shape the format of a particular record

type. Some of the factors operate independently, whereas others pull in opposing directions, meaning that high affordance is given to one at the expense of the other. This is of course a balancing act, and though it is one that is often necessary, it has the potential to result in compromised evidential value if it is not reflected upon and subjected to careful scrutiny. There is nothing inherently wrong with written records of this kind being shaped by a series of trade-offs and cost/benefit calculations, since clearly they are being produced in a context of limited resources and in which the records necessarily get repurposed throughout their journey through the criminal justice system, but these calculations must be acknowledged as such, and it must be ensured that the principles underlying them are sound and justifiable.

Making these factors and trade-offs more visible enables us to see clearly why, from our perspective, police interview record formats are often seemingly not fit for their intended purpose. It enables us to take account of the practical realities of the context, including the very real problems of lack of time and money, but still show how a better balance can be struck, which recognizes and accepts those unavoidable limitations, but nevertheless can point to how this can best be compensated for. This approach also highlights that one of the most effective targets for improvement may be in increasing efficiency in the production of transcripts, and that is one aspect of the research we are currently undertaking (Deamer et al., in review). However, this still needs to be balanced against the practical requirements of the context, where longer or more detailed transcripts may result in an unacceptable cost in terms of usability and fitness for purpose.

We do not propose this as a complete solution, instead we propose this as a starting point for a new way to approach the question of transcript adequacy. In particular we recognize the challenge that identifying one specific purpose is itself not always straightforward, especially in legal contexts. Further, while we envisage that this is useful and useable across a wide range of contexts, it has of course been developed with one quite specific context in mind (police interview records); we therefore hope that others will be willing to apply this to their own contexts, either academic or practitioner, to see whether it is sufficiently universal or perhaps requires further adaptation, either to become genuinely universal, or perhaps needing slightly different approaches for particular contexts.

Finally, we reiterate encouraging reflective consideration of our own practices as researchers as we transcribe, analyse, and represent spoken data. As linguists, we too have a power over those whose talk we transcribe, and come to the process of transcription with our own 'professional hearings' and with our own institutional goals (Bucholtz 2009). In particular, the approach we propose here prompts questions such as to whether we could (or should) give more agency to our participants, including more control over the written version of their words which we produce for public scrutiny.

In 2014, Fraser stated a first step in solving the general problems with reliability of transcription was to publicize them. We hope to have contributed to

that here, and hope that this both prompts further research on institutional transcription, and reinvigorates an important methodological debate in applied linguistics.

Conflict of interest statement. None declared.

NOTES

1 *Source*: https://webarchive.nationalarchives.gov.uk/20090808155206/http://www.the-shipman-inquiry.org.uk/trialtrans.asp (accessed 30 June 2021).

2 Procedure in England and Wales for challenging records of interview, requesting that a recording is played in court, and related matters, is set out in the Criminal Practice Directions Amendment No 3 [2015] EWCA Crim 430, para 27C.

3 From Haworth's own experience as a barrister, it is relatively common in England and Wales for Prosecution and Defence to agree edits to the version presented to court on other grounds, such as legal admissibility or prejudice.

REFERENCES

Arnold, L. 2021. 'Strategies used by suspects during a police interview', unpublished PhD thesis, University of Portsmouth.

Ashmore, M., **K. MacMillan**, and **S. D. Brown**. 2004. '"It's a scream": Professional hearing and tape fetishism,' *Journal of Pragmatics* 36: 349–74.

Auburn, T., **S. Drake**, and **C. Willig**. 1995. '"You punched him, didn't you?": Versions of violence in accusatory interviews,' *Discourse & Society* 6: 353–86.

Bauman, R. and **C. L. Briggs**. 1990. 'Poetics and performance as critical perspectives on language and social life,' *Annual Review of Anthropology* 19: 59–88.

Biber, D. 1988. *Variation across Speech and Writing*. Cambridge University Press.

Blackwell, S. 1996. 'Corrective measures: Some aspects of transcription in the British legal system' in K. Hannes (ed.): *Recent Developments in Forensic Linguistics*. Peter Lang.

Blommaert, J. 2001. 'Investigating narrative inequality: African asylum seekers' stories in Belgium,' *Discourse & Society* 12: 413–49.

Blommaert, J. 2005. *Discourse*. Cambridge University Press.

Bucholtz, M. 2000. 'The politics of transcription,' *Journal of Pragmatics* 32: 1439–65.

Bucholtz, M. 2007. 'Variation in transcription,' *Discourse Studies* 9: 784–808.

Bucholtz, M. 2009. 'Captured on tape: Professional hearing and competing entextualizations in the criminal justice system,' *Text & Talk* 29: 503–23.

Bucke, T., **R. Street**, and **D. Brown**. 2000. 'The right of silence: The impact of the Criminal Justice and Public Order Act 1994,' *Home Office Research Study 199*. Home Office. https://www.worldcat.org/title/right-of-silence-the-impact-of-the-criminal-justice-and-public-order-act-1994/oclc/450358215?referer=di&ht=edition.

Canning, P. 2018. '"No ordinary crowd": Foregrounding a "hooligan schema" in the construction of witness narratives following the Hillsborough football stadium disaster,' *Discourse & Society* 29: 237–55.

Canning, P. 2020. 'Retelling Hillsborough: A critical stylistics analysis of witness statements' in M. Lambrou (ed.): *Narrative Retellings: Stylistic Approaches*. Bloomsbury.

Cook, G. 1990. 'Transcribing infinity: Problems of context presentation,' *Journal of Pragmatics* 14: 1–24.

Cotterill, J. 2003. *Language and Power in Court: A Linguistic Analysis of the O.J. Simpson Trial*. Palgrave.

Coulthard, M. 1996. 'The official version: Audience manipulation in police records of interviews with suspects' in C. Caldas-Coulthard and M. Coulthard (eds): *Texts and Practices: Readings in Critical Discourse Analysis*. Routledge.

Coulthard, M. 2002. 'Whose voice is it? Invented and concealed dialogue in written records of

verbal evidence produced by the police' in J. Cotterill (ed.): *Language in the Legal Process*. Palgrave Macmillan.

Cribb, M. and **S. Rochford**. 2017. 'The transcription and representation of spoken political discourse in the UK House of Commons,' *International Journal of English Linguistics* 8: 1.

Deamer, F., **E. Richardson**, and **K. Haworth**. In review. 'For the Record: Exploring variability in interpretations of police investigative interviews,' *Language and Law/Linguagem e Direit*.

Dixon, D. and **G. Travis**. 2007. *Interrogating Images: Audio-Visually Recorded Police Questioning of Suspects*. The Sydney Institute of Criminology.

Eades, D. 1996. 'Verbatim courtroom transcripts and discourse analysis' in H. Kniffka, S. Blackwell and M. Coulthard (eds): *Recent Developments in Forensic Linguistics*. Peter Lang.

Fraser, H. 2003. 'Issues in transcription: Factors affecting the reliability of transcripts as evidence in legal cases,' *International Journal of Speech, Language and the Law—Forensic Linguistics* 10: 203–26.

Fraser, H. 2014. 'Transcription of indistinct forensic recordings,' *Language and Law/Linguagem e Direito* 1: 5–21.

Gibbons, J. 1995. 'What got lost? The place of electronic recording and interpreters in police interviews' in D. Eades (ed.): *Language in Evidence: Linguistic and Legal Perspectives in Multicultural Australia*. University of New South Wales Press.

Gibbons, J. 2001. 'Legal transformations in Spanish: An "audiencia" in Chile,' *Forensic Linguistics* 8: 24–43.

Gibbons, J. 2003. *Forensic Linguistics: An Introduction to Language in the Justice System* Blackwell.

Halliday, M. A. K. 1989. *Spoken and Written Language*. Oxford University Press.

Haworth, K. 2009. 'An analysis of police interview discourse and its role(s) in the judicial process,' unpublished PhD thesis, University of Nottingham.

Haworth, K. 2013. 'Audience design in the police interview: The interactional and judicial consequences of audience orientation,' *Language in Society* 42: 45–69.

Haworth, K. 2017. 'The discursive construction of evidence in police interviews: Case study of a rape suspect,' *Applied Linguistics* 38: 194–214.

Haworth, K. 2018. 'Tapes, transcripts and trials: The routine contamination of police interview evidence,' *The International Journal of Evidence & Proof* 22: 428–50.

Haworth, K. 2020. 'Police interviews as evidence' in M. Coulthard, A. May and R. Sousa-Silva (eds): *Routledge Handbook of Forensic Linguistics*. Routledge.

Heydon, G. 2005. *The Language of Police Interviewing: A Critical Analysis*. Palgrave Macmillan.

Jefferson, G. 2004. 'Glossary of transcript symbols with an introduction' in G. Lerner (ed.): *Conversation Analysis: Studies from the First Generation*. John Benjamins.

Jefferson, G. 1985. 'An exercise in the transcription and analysis of laughter'in T. Van Dijk (ed.): *Handbook of Discourse Analysis, Vol. 3: Discourse and Dialogue*. Academic Press.

Jefferson, G. 1996. 'A case of transcriptional stereotyping,' *Journal of Pragmatics* 26: 159–70.

Jönsson, L. and **P. E. R. Linell**. 1991. 'Story generations: From dialogical Interviews to written reports in police interrogations,' *Text—Interdisciplinary Journal for the Study of Discourse* 11: 419.

Komter, M. L. 2012. 'The career of a suspect's statement: Talk, text, context,' *Discourse Studies* 14: 731–52.

Komter, M. L. 2019. *The Suspect's Statement: Talk and Text in the Criminal Process*. Cambridge University Press.

Lapadat, J. C. 2000. 'Problematizing transcription: Purpose, paradigm and quality,' *International Journal of Social Research Methodology* 3: 203–19.

Linell, P. E. R. 1998. 'Discourse across boundaries: On recontextualizations and the blending of voices in Professional discourse,' *Text—Interdisciplinary Journal for the Study of Discourse* 18: 143.

MacLeod, N. 2020. '"Tell me in your own words . . .": Reconciling institutional salience and witness-compatible language in police interviews with women reporting rape' in M. Mason and F. Rock (eds): *The Discourse of Police Interviews*. University of Chicago Press.

Maybin, J. 2017. 'Textual trajectories: Theoretical roots and institutional consequences,' *Text & Talk* 37: 415–35.

Meredith, J. 2016. 'Transcribing screen-capture data: The process of developing a transcription system for multi-modal text-based data,' *International Journal of Social Research Methodology* 19: 663–76.

Mollin, S. 2007. 'The Hansard hazard: Gauging the accuracy of British parliamentary transcripts,' *Corpora* 2: 187–210.

Mondada, L. 2007. 'Commentary: Transcript variations and the indexicality of transcribing practices,' *Discourse Studies* 9: 809/809–21.

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

Park, J. S.-Y. and **M. Bucholtz**. 2009. 'Introduction. Public transcripts: Entextualization and linguistic representation in institutional contexts,' *Text & Talk* 29: 485–502.

Rock, F. 2001. 'The genesis of a witness statement,' *Forensic Linguistics* 8: 44–72.

Sacks, H., **E. A. Schegloff**, and **G. Jefferson**. 1974. 'A simplest systematics for the organization of turn-taking for conversation,' *Language* 50: 696–735.

Sarangi, S. 1998. 'Rethinking recontextualization in Professional discourse studies: An epilogue,' *Text* 18: 301–18.

Shipman, R. V. (Harold Frederick), 1999. Unreported.

Shuy, R. W. 1993. *Language Crimes: The Use and Abuse of Language Evidence in the Courtroom*. Blackwell.

Slembrouck, S. 1992. 'The parliamentary Hansard "verbatim" report: The written construction of spoken discourse,' *Language and Literature: International Journal of Stylistics* 1: 101–19.

Stokoe, E. 2009. '"For the benefit of the tape": Formulating embodied conduct in designedly uni-modal recorded police–suspect interrogations,' *Journal of Pragmatics* 41: 1887–904.

Walker, A. G. 1990. 'Language at work in the law: The customs, conventions, and appellate consequences of court reporting' in J. N. Levi & A. G. Walker (eds.): *Language in the Judicial Process*. Plenum Press.

NOTES ON CONTRIBUTOR

Dr Emma Richardson is a Research Fellow within the Aston Institute for Forensic Linguistics. Her research interests centre on improving access to justice for vulnerable victims and witnesses of crimes such as domestic violence and sexual assault offence. She is a conversation analyst interested in how written policy and guidance documents translated in spoken talk and used by people, in training and in practice across forensic and health care settings.

Dr Kate Haworth is Director of the centre for research on Spoken Interaction in Legal Contexts (SILC), Institute for Forensic Linguistics, Aston University. Her research interests include all aspects of language and communication in legal contexts, especially the use of language data as evidence. Drawing on her own professional background as a barrister, she is keen to strengthen links between academic research and professional practice.

Dr Felicity Deamer is a Research Fellow based in the Aston Institute for Forensic Linguistics. Her research occupies the intersection between Forensic Linguistics, Pragmatics, and Language and Mental Health. She uses a variety of methods, from conversation analysis to eye-tracking, to investigate a variety of phenomena, including police investigative interviews, violent and abusing interactions on mental health wards, patient/therapist interactions, and metaphor comprehension in psychosis.