Forensic transcription:
How confident false beliefs about language and speech threaten the right to a fair trial in Australia

Helen Fraser
Adjunct Associate Professor, University of New England
helenbfraser@gmail.com
To appear in Australian Journal of Linguistics, Nov 2018

Abstract

Everyday knowledge about language and speech, or ‘folk linguistics’, incorporates a number of false beliefs that have a negative effect in a range of areas, nowhere more so than in the criminal justice system (e.g. Solan, Ainsworth & Shuy, 2015). One lesser known area where false beliefs have a major impact is forensic transcription (interpretation of indistinct covert recordings used as evidence in criminal trials). Without consultation of the linguistic sciences, the law has developed processes that allow police transcripts to ‘assist’ the courts in making out unintelligible covert recordings (Fraser 2017a). The present article uses a case study of a real murder trial to bring the actual and potential injustice this creates to the attention of linguistic science, and examine the issues from a linguistics perspective. Having laid out the problem, it goes on to consider potential solutions, arguing that creating a better process requires deep collaboration between linguistics, law and law enforcement. From linguists, it requires improved theoretical understanding of the nature and structure of the false beliefs that underlie the existing legal processes, as well as development of a more general theoretical account of the process of transcription and the nature of transcripts.

1. Common knowledge versus expertise in linguistics

Linguists are familiar with the fact that mastering our field requires, before learning new information about language and speech, unlearning false beliefs widely accepted by society as ‘common knowledge’ (e.g. Bauer & Trudgill, 1998; Niedzielski & Preston, 2003). Most linguists, having achieved sufficient unlearning, understandably prefer to leave false beliefs behind and focus on generating new, reliable knowledge. Unfortunately, however, only a small proportion of that reliable knowledge permeates the membrane between academia and the wider community, where many false beliefs about language and speech continue to flourish, creating problems in a variety of domains (e.g. Grant, 2014; Lippi-Green, 2012). As the saying goes, it’s not what people don’t know that causes trouble, but what they know for sure that’s just not so.
Perhaps the most troublesome of all false beliefs are those held by people who are highly educated in disciplines other than linguistics. Such individuals often have strong confidence in what we might call educated common knowledge about language, which they may contrast favourably with uneducated common knowledge, not realising that their own more sophisticated beliefs are no less false from the point of view of the linguistic sciences. One area where this is particularly true is the law. Despite some successful education about the specialised knowledge contributed by linguistics (e.g. Eades, 2010), many aspects of language and speech are still considered matters of (educated) common knowledge, about which lawyers can pronounce with confidence.

This is significant, since expert opinion is only sought by the law about matters of specialised knowledge. For evidence whose interpretation is considered to require only common knowledge, readily available to any jury, not only is it unnecessary to call upon an expert, doing so is actively discouraged (cf. Allsop, 2016). Such lack of consultation allows false beliefs about language and speech to be given unwarranted credence among lawyers, as is well documented for multiple areas of law by the field of forensic linguistics (e.g. Coulthard, Johnson, & Wright, 2017; Solan, Ainsworth, & Shuy, 2015). The corollary is that, unconsulted, linguists remain unaware of legal practices involving language and speech evidence that cause real problems in the criminal justice process.

One example is the handling of indistinct covert recordings used as evidence in court, where several practices commonly used in our legal process threaten the right to a fair trial (Fraser, 2018b). The present paper aims to bring one of these practices (admission of police transcripts as ‘assistance’ to the jury in interpreting indistinct audio) to the attention of linguists who may be willing and able to help bring about appropriate reform.

Two points should be noted briefly before starting. First, though many of the issues raised are also relevant internationally, the focus of the current paper is on the Australian context. Second, this is a high-level treatment of complex matters detailed in publications spanning fifteen years. For those unfamiliar with this background, the succinct overviews provided by Burridge (2017) or Fraser (2017b) are strongly recommended. More general background, including audio demonstrations, is available at forensictranscription.com.au.

2. Disputed utterances but undisputed transcripts

Covert recordings (captured via telephone intercept, ‘bugging’ of homes or cars, body wires worn by undercover agents, etc.) are routinely collected during investigation of major crimes, providing vital intelligence to investigators. Many are later admitted in trials, where they can offer powerful evidence, allowing the court to hear speakers making admissions they would not be prepared to give overtly. One difficulty is that, since it is hard to control the recording conditions, the audio is often of extremely poor quality, to the extent of being unintelligible without assistance.
During the 1980s, lawyers observed that detectives can often make out more than other listeners in indistinct recordings related to their cases. Further, transcripts prepared by these detectives can assist the perception of other listeners. Following a landmark High Court decision (Butera v DPP (1987) 164 CLR 180), it is now routine for indistinct covert recordings admitted as evidence in court to be accompanied by transcripts intended to assist the jury in interpreting the audio, and for the transcripts to be produced by police given the role of so-called ‘ad hoc expert’ (more on this below; for background see Fraser (2015; 2017a)).

Of course the risk is recognised that the jury might simply read a (potentially misleading) police transcript and accept it as a reliable representation of the recording. To ensure this does not happen, the law has implemented a range of precautionary practices. These are complex but rest on two main tenets. The first tenet is the expectation that the defence will check the prosecution’s transcript thoroughly. Where they disagree with any part, they should bring their own interpretation to the attention of the prosecution, with the aim of creating an agreed version. If agreement is not possible, either or both parties can seek assistance from an expert (often an audio engineer, but increasingly, thanks to ongoing education of lawyers, a qualified phonetician).

This is the part of the legal process that has given rise to the branch of forensic phonetics called disputed utterance resolution (e.g. French & Stevens, 2013). A number of cases have become famous via phonetics experts revealing crucial errors in prosecution transcripts. For one oft-cited example, I shot a man to kill was shown to be I showed a man a ticket (e.g. Coulthard et al., 2017, p. 132).

The existence of this branch understandably gives other linguists the impression that the topic of forensic transcription is appropriately taken care of. That is far from true, as the present paper seeks to demonstrate. In fact, defence teams rarely dispute a police transcript, and even more rarely seek expert assistance, as discussed further below.

With or without dispute, audio and transcripts are evaluated during the voir dire (a pre-trial hearing to determine the admissibility of evidence). In case of dispute, the judge may review the material personally, and can, in principle, rule one or more transcripts, or even the audio itself, to be inadmissible. This is rare though. Far more commonly, competing transcripts are provided to the jury, for them to decide which they find most helpful in interpreting the audio. Phoneticians who provide expert

1 Unfortunately, there are few resources to enable proper quantification of cases involving covert recordings of various kinds used in various ways (see Fraser, 2018). The rarity of expert consultation can be seen informally by comparing the relatively small discipline of forensic phonetics with the enormous quantity of covert audio generated during investigations, and noting that, even within forensic phonetics, disputed utterance resolution is a minor branch: far more expert evidence is sought about speaker identification (French & Stevens, 2013).
opinions about disputed utterances may thus be surprised to learn that their views are admitted, not instead of, but as an optional alternative to, the police opinion.

All this is explained by the second tenet: the concept that the evidence is the audio, not the transcript. On the view that understanding English language requires only common knowledge, interpretation of covert recordings is considered a matter for the jury. The transcript is deemed a mere convenience, offering assistance much as a chart assists in interpreting an array of figures. To ensure the jury understands this, the judge is required to instruct them carefully, using directions modelled on standard versions provided in bench books, that a transcript is just one person’s opinion (rather strangely called an ‘aide memoire’ - see Fraser, 2015 for background), which jurors may or may not find helpful. It is essential for them to listen carefully to the audio and reach their own opinion as to the contents of the recording.

3. A bad situation

As will be unsurprising to many readers, the process described above is not always effective in ensuring reliable interpretation of indistinct covert recordings. This is for two sets of reasons.

First, creating a reliable and useful transcript of a forensic recording is a complex task requiring real, not ‘ad hoc’, expertise (for a full critique of the highly problematic concept of ‘ad hoc expert’, see Edmond and San Roque, 2009). Even with high quality recordings, transcription of conversational speech raises many theoretical and practical issues regarding representation and layout. While some of these are well known to disciplines such as phonetics and conversation analysis (e.g. Shockey, 2003; Sidnell & Stivers, 2012), the forensic context creates additional complexity (Fraser, 2014). Indistinct audio is even more problematic, requiring expert management of contextual priming – the tendency for a transcriber to be unconsciously influenced by prior expectations or assumptions about the content of the recording (Fraser, 2003).

Police, however, receive no training of any kind in how to do forensic transcription. This makes it unlikely their transcripts would provide a fully reliable representation of indistinct covert recordings. Indeed, in my experience, they typically show substantial problems in both content and layout. A brief impression can be gained from Example 1 (discussed further below).

Example 1

Note: ‘ind’ and ‘.’, represent ‘indistinct’ sections of unspecified duration.

Play this game ..[ind].. [Track 17]..[ind].. play the ..[ind]..game ..[ind].. Just play the fuckin ..[ind].

Second, with indistinct audio, reliably evaluating another person’s transcript requires even greater expertise than creating one’s own, due to the effect of priming by a transcript. This is an extraordinary phenomenon, whose power is hard to appreciate by those who have not personally experienced it (again Fraser (2017b) is
recommended as a succinct introduction). Recent experiments (Fraser, 2018c; Fraser & Stevenson, 2014; Fraser, Stevenson, & Marks, 2011) demonstrate how seeing a transcript can prime listeners to confidently ‘hear’ suggested words in audio they previously found unintelligible – even if those words are manifestly inaccurate. This means a transcript can be experienced as ‘assistance’, even when it is actually misleading. Most importantly, having once ‘heard’ suggested words, listeners are unlikely to fully ‘unhear’ them when advised they are factually wrong. This gives the inaccurate transcript a ‘continued influence’ (Ecker, Lewandowsky, & Tang, 2010) even on those who come to disagree with some parts. The experiments also show that priming by a transcript is especially powerful in conjunction with contextual priming, and can have a significant effect on participants ‘verdicts’ of guilt.

These two sets of factors make it extremely difficult for jurors to reach an independent interpretation of indistinct audio under courtroom conditions, with barristers on either side exhorting them to consider their interpretation (Fraser, 2013).

Of course we can never know for sure what conclusion a real jury reaches regarding the content of forensic audio. However I am personally aware of numerous cases, one of which forms the case study below, where manifestly inaccurate police transcripts have been admitted as ‘assistance’ in making out indistinct covert recordings, resulting in actual or potential injustice.

4. Confident False Beliefs Of Educated Common Knowledge

Linguists told of this situation sometimes take a cynical view, suggesting the legal process has been developed to benefit prosecutors. My understanding is that the process was developed with the honest intention of assisting juries to reach a fair verdict. The problem is simply that it has been created, modified and administered over a thirty-year period with no consultation of the linguistic sciences, giving full rein to educated common knowledge about language and speech (Fraser, 2018b). The result is a baroque structure of false beliefs so confidently accepted that, far from yielding to expert opinion, they are actually used as the basis of evaluating the opinions of genuine experts (examples below).

Some of these false beliefs go beyond those commonly encountered by linguists via ‘folk linguistics’ (though see Coulthard, 1997; Shuy, 2007; Tiersma, 2009 for examples of similar issues from related branches of forensic linguistics). Indeed, it can be hard for linguists to believe that anyone could seriously hold such views.

Interestingly, for lawyers who have become used to their way of doing things, it seems equally difficult to accept the linguists’ perspective (see further discussion below). This clearly creates a major barrier to the interdisciplinary communication required for progress (cf. Fraser & Schalley, 2011).

The question is – how to overcome this barrier. Well-established principles of human learning (e.g. National Research Council, 2005 Ch.9) suggest it is not enough for
linguists to provide information about language and speech to lawyers, since laying new information on top of false beliefs often results in imperfect understanding. In order to educate successfully, it is necessary to discover the nature of the false beliefs, and develop evidence-based strategies for countering them effectively.

This section attempts a preliminary exegesis of some key false beliefs about language and speech that underpin the legal handling of covert recordings. Ideally it will soon be supported by a survey of lawyers’ attitudes. In the meantime it is based on extensive personal experience, exemplified by the case study that follows. As noted, several of the ideas find resonance in the literature on folk linguistics and language ideology.

4.1. False Belief: Expertise In Transcription Comes From Listening Many Times

In general, police are restricted to giving factual evidence in court. Opinion evidence is accepted only from expert witnesses, i.e. those recognised as having ‘specialised knowledge based on training, study or experience’ (Evidence Act Section 79). However the process outlined above requires detectives’ transcripts to be provided as their personal opinion, for evaluation by the jury. To enable this, police transcribers are given the status of ad hoc expert, on grounds that they have specialised knowledge in relation to the specific recording they have transcribed. This requires explaining their apparent advantage in transcription as deriving from some form of ‘expertise’ – which in this case is attributed to their having listened to the recording ‘many times’.

To linguists with genuine expertise in transcribing indistinct conversational speech, this is clearly the wrong explanation. Listening many times is certainly necessary, but far from sufficient: it is quite possible to listen many times and be wrong every time.

Whatever advantage detectives may appear to have in interpreting indistinct covert recordings related to their cases, it comes not from listening many times, but from their contextual knowledge about the circumstances of the case. And it is important to recognise that this is indeed a valuable source of information, that should be harnessed appropriately in order to achieve reliable interpretation of indistinct audio (see Fraser, 2015). However, contextual knowledge in itself confers no expertise in transcription. Indeed, it is a double-edged sword, since the role of contextual priming makes it just as easy for unreliable assumptions to mislead, as it is for reliable knowledge to assist, with, surprising as it may seem, nothing in the perceptual experience of the listener to differentiate these two scenarios (Fraser, 2003; 2014).

Lawyers however, seem to really believe the ‘listening many times’ explanation. For example, when phonetic analysis has caused me to disagree with a police transcript that has previously been accepted by both sides, even defence lawyers (to whose client the opinion I am offering would be an advantage) have suggested, kindly and apparently in good faith, that maybe I should try listening a few more times.
Clearly these lawyers have accepted the police transcript to a degree that resists challenge by an alternative opinion, even from someone they respect as an expert. The reason is suggested by the next false belief.

4.2. *False belief: A transcript is verified by the extent to which it assists the perception of responsible listeners*

Defence lawyers typically read a police transcript as part of the prosecution brief. Of course they then check the transcript against the audio. What they may not realise, however, is that ‘checking’ is a surprisingly ineffective means of evaluating a transcript. Even with a clear recording and a professional transcript, it is a painstaking and time-consuming task, liable to leave many errors unnoticed (consider for example the sometimes entertaining mistakes found in transcripts of radio broadcasts).

With poor quality audio, the difficulty is far greater, as the audio rushes by quickly, making it hard to pay attention to every phrase. In the context of trust that police who have listened many times have probably got most of it right, defence teams’ recognition of a few key phrases in the audio is enough to create credibility for the police transcript as a whole (Fraser, 2013). This apparent credibility is greatly enhanced if some of those key phrases seem to be supported by external evidence, since listeners tend to disregard the degree to which the detective’s knowledge of the external context assisted in producing the transcript in the first place – creating a dangerous circularity of reasoning (Fraser 2017a).

This problem is greatly exacerbated by the unprofessional layout typical of many police transcripts. Consider Example 1 above. With no indication of the nature or duration of ‘indistinct’ sections, listeners quickly get lost. Minimal acoustic cues are sufficient to confirm the suggested words ‘play’ and ‘game’, and make plausible a general scenario, heavily but misleadingly primed by the transcript, in which one speaker repeatedly urges the other to ‘play the game’.

Of course, defence lawyers do not rely only on their own hearing. They expect the defendant to check the police transcript as well. Though this may seem to offer the ultimate verification, it too is fraught with false beliefs.

4.3. *False belief: A defendant can readily dispute mistranscribed utterances in an indistinct covert recording*

The legal process outlined above rests strongly on the understanding that the defendant, who presumably knows what was really said in the recorded conversation, has a fair opportunity to dispute incriminating sections of the police transcript. However, it is not enough just to deny that particular words were spoken. To demonstrate convincingly that the police version has been inaccurately transcribed, the defendant is expected to provide a plausible alternative version. This idea incorporates several problems.
First, it compromises the basic principle of law that it is up to prosecutors to prove guilt, not the defendant to prove innocence (see further comment in Section 4.4).

Second, it can be surprisingly hard, even when the police transcript is factually inaccurate, for an innocent defendant to provide exact words as an alternative. While participants in a conversation can be expected to recall the gist of what was said, memory for specific utterances is known to be very poor (e.g. Davis & Friedman, 2007). Of course, defendants have the recording to jog their memory, but this offers less help than lawyers imagine. As some readers may attest from their own experience, with very indistinct audio it can be genuinely difficult, even for a participant in the original conversation, to write out individual utterances in detail – and of course doing so requires literacy and cognitive skills not in the reach of every defendant.

Under these conditions, a police transcript can prime a participant’s perception as easily as anyone else’s. Thus it is not uncommon for a defendant who denies having said words in a police transcript to allege the audio has been tampered with, not recognising the real explanation is far simpler: inaccurate transcription.

Finally, even if the defendant does suggest an alternative, it is still necessary to persuade defence lawyers to challenge the prosecution version. Whether or not they personally agree with the defendant’s interpretation, recognising the possibility the jury might prefer the police version might cause them to advise building a defence on the basis of an alternative interpretation of the words, rather than on disputing the words themselves.

4.4. Ultimate false belief: Speech perception is picking up ‘what is there to be heard’

The above false beliefs are not a random collection of unrelated concepts, but part of a complex network arising from more basic ideas about the nature of speech, and how it is processed. An important component of educated common knowledge about language (e.g. Linell, 2005; Olson, 1994) is the idea that speech is similar to print – a sequence of discrete, invariant phonemes grouped into words, phrases and sentences. According to this ‘literacy bias’ (Fraser, 2010), speech perception is a matter of picking up the sound patterns corresponding to individual phonemes, putting them together into words and sentences, and then interpreting their meaning in context – much as a beginning reader ‘sounds out’ letters on a printed page.

Of course, this is nothing like how speech perception really works. Speech is a continuous stream of sound (e.g. Ladefoged & Disner, 2012). Dividing it into words, phonemes and other units requires complex, though unconscious, cognitive processing. Importantly, recognition of apparently basic units, like phonemes, depends to a surprising degree upon prior interpretation of words and sentences, which in turn depends to a surprising degree upon prior understanding of meaning and context (e.g. Cutler, 2012; Warren, 2012). While this is well established by the phonetic sciences, the literacy bias exerts a powerful influence on educated common
knowledge – indeed some argue remnants even affect linguistic theories (e.g. R. Harris, 1981; Magnuson, Mirman, & Myers, 2013; Port & Leary, 2005).

In particular, the literacy bias is deeply entrenched in the law, to the extent that covert recordings are explicitly treated as ‘documents’ (e.g. Evidence Act Section 48), with many other ideas built on this infelicitous analogy (see further discussion in Fraser, 2015). One of these ideas is that transcription is a straightforward matter of writing down words that are objectively ‘there to be heard’ (see example in Innes, 2011, p. 151). Another idea, relevant to the case study below, is that poor quality audio can be improved by ‘enhancing’, creating a supposedly ‘clearer’ version of the ‘document’, which may be admitted with minimal scrutiny (Fraser, 2018b). Actually, typical ‘enhancing’ techniques rarely make any objective improvement to intelligibility, though listeners’ belief they make the audio ‘clearer’ can enhance the credibility of an inaccurate transcript (see Fraser 2015).

Most problematic of all is the idea that a transcript offers optional ‘assistance’ to a jury listening to indistinct audio, when in fact it strongly primes their perception (validly or otherwise). It is important to be clear that this does not mean lawyers fail to recognise the potential of an inaccurate transcript to influence a jury’s hearing. Indeed the law is powerfully aware of the danger of suggestibility, and includes numerous measures to ensure that prejudicial evidence is kept away from the jury. The problem is that the legal process outlined above, designed to protect the jury from being misled by an inaccurate transcript, depends on lawyers acting as the gatekeepers who exclude misleading transcripts before they can influence the jury. This assumes lawyers’ own perception is veridical. Indeed lawyers seem to consider themselves to be immune from being misled by an inaccurate transcript – perhaps a manifestation of the ‘bias blind spot’ that affects us all, making us far more likely to notice others’ biases than our own (Pronin, Lin, & Ross, 2002).

Actually, no one is immune from priming, even linguists (Fraser, 2014). Linguists who might doubt their own susceptibility are urged to read Wald (1995). Avoidance of priming can only be achieved at a system level, not via the will of individuals (Fraser 2013; Thompson 2009). This is why it is so problematic to start from the police interpretation of indistinct covert recordings. No matter how strongly subsequent listeners are directed to use it only as assistance, the power of priming means the police version inevitably gets perceptual privilege, to the extent of compromising the principle ‘innocent till proven guilty’ – as exemplified in the case study.

5. Case study

This section aims to give an impression of how false beliefs such as those discussed above play out in real trials. Before starting, I want to emphasise that, though clearly I am critiquing specific anomalies in the legal process, I do so out of respect both for the institution of the law and for those administering it. Indeed, while accountability for failings in the current legal process for handling indistinct covert recordings surely

Fraser: Forensic Transcription

9
lies with the judiciary (Fraser, 2018b), arguably some responsibility for the continuing existence of the false beliefs that support the process must be taken by linguists: as a discipline, we have evidently failed to disseminate the findings of our science to all relevant quarters.

5.1. The trial

The case study involves a very complex murder trial, here selectively simplified to highlight aspects relevant to present themes (more background is available in Fraser (2017a)).

A major piece of evidence was an extremely indistinct covert recording, transcribed by a detective as including the phrase ‘at the start we made a pact’ (the ‘pact’ phrase). The key issue in the trial was whether it was a pact to commit the murder (making the defendant party to a joint criminal enterprise, as guilty of murder as if he had personally pulled the trigger) or a pact to conceal the murder (making him an accessory after the fact – a serious crime, but not nearly as serious as murder). Importantly to the case study, the defendant admitted being an accessory after the fact. Ultimately, however, the jury reached a verdict of murder and a sentence of thirty years was imposed. Several unsuccessful attempts at appeal were made, including one to the High Court, none having anything to do with the audio or transcript.

About a year later, I was asked to examine the audio (in general; the ‘pact’ phrase was not specifically mentioned). After considerable investigation, I determined that the police transcript was inaccurate and misleading throughout. Looking, later, at the trial transcripts, I noted how much emphasis was given to the phrase ‘at the start we made a pact’, and returned to analyse that section in more detail. I concluded that the ‘pact’ phrase was not just inaccurate but phonetically implausible. Since by then I was already deeply concerned about the legal process for handling indistinct covert recordings, I used the audio for an experiment (Fraser & Stevenson, 2014) aiming to demonstrate how easy it is, even when the legal processes outlined above are followed perfectly (as they evidently were in this trial), for a manifestly unreliable transcript to be admitted as ‘assistance’ to a jury; and how likely it is that a jury, though earnestly following the judge’s direction to listen carefully and form their own opinion, will nevertheless be unwittingly misled by it.

Briefly, the experiment showed that, of 138 participants, none heard even the word ‘pact’, let alone anything remotely like the ‘pact’ phrase, unless it was first suggested to them. Even when the phrase was suggested, in the absence of background context it showed an unusually weak priming effect, and was readily abandoned for an alternative phrase. This alternative phrase, ‘it’s fucking payback’, was made up for the purpose of the experiment, with emphasis that it was not posited as an accurate transcript, merely as a phonetically more plausible interpretation – noting that substantial differences in both segmental and suprasegmental structure between the two phrases confirm the implausibility of the ‘pact’ phrase.
However, despite its objective implausibility, when ‘at the start we made a pact’ was suggested in the context of the story of the murder, it had a strong priming effect, was far less readily abandoned, and significantly influenced participants’ ‘verdicts’ of guilt.

5.2. The application for review of the conviction

While doing this work, I observed (from the trial transcripts) that the ‘circumstantial’ evidence in the case only gained significance in light of the general concept that there had been a pact of some kind. However, the only ‘direct’ evidence for any pact at all was the detective’s inaccurate transcription of the ‘pact’ phrase. Nevertheless, acceptance of the ‘pact’ phrase seemed to have created a kind of tunnel vision strongly affecting interpretation of the circumstantial evidence – a known risk for wrongful conviction (Gould, Carrano, Leo, & Young, 2012).

Of course I could not judge whether or not the convicted murderer was factually innocent, but I became convinced he had not had a fair trial. His lawyers advised that, since all opportunities for appeal had been exhausted on other grounds, the only avenue remaining was a review of the conviction. This is a complex process in which successful outcomes are extremely rare. The main requirement is ‘fresh evidence’ (i.e. not available at the time of the original trial) which could have made a difference to the original verdict. The first step is acceptance of an application to present the case for review, which depends on a single judge being given ‘a sense of unease or disquiet’ regarding the original conviction.

Accordingly I put together a lengthy report which included: detailed analysis of the audio (which existed in multiple poorly documented versions); demonstration that the police transcript (also in multiple poorly documented versions) is unreliable and misleading throughout; detailed phonetic analysis of the pact section, concluding ‘I am confident it is not at the start we made a pact but I don’t know what it is’; and results (summarised above) of Fraser and Stevenson (2014). This report was eventually submitted as part of an application for a review of the conviction, which then required a response from the Crown, and a rejoinder from the defence, before ultimately receiving a decision from the judge as to whether the application was successful – in which case, the actual review process could have begun.

The application was rejected. Though disappointing, this was unsurprising, given the low rate of successful outcomes. What was surprising was the reasons given for the rejection, which demonstrated substantial misinformation about language and speech. The following sections briefly summarise a few key points from the Crown response, the judge’s decision, and the defence submissions.

5.3. The Crown response

I had been given to understand the Crown respondent (a barrister with the Office of the Director of Public Prosecutions) would send my report to another independent expert for review. In fact he reviewed it himself, concluding that my report was not
fresh evidence and would have made no difference to the outcome of the trial. My opinion was merely that of another person who had listened many times, and would not have assisted the court any more than that of the detective – especially as I had not had the benefit of the infrared headphones provided to the jury (‘infrared’, of course, indicates only that the headphones used in court were wireless; my wired headphones were undoubtedly of superior quality). He also noted a ‘logical difficulty’ in my asserting both that the audio was too indistinct for me to state what was said, and that the ‘pact’ phrase was inaccurate. Personally, I see no contradiction in these assertions, but this view (‘If you don't know what the words are you can't say what they are not’) is commonly expressed by lawyers.

The main burden of his argument was that, since the audio and transcript had been admitted in accordance with legal authority, it was fully open to the jury to reach their own opinion regarding the contents of the covert recording. This of course is true – but sidesteps the question of whether it was open to them to reach the correct opinion: my report of course had argued strongly that it was not.

The Crown respondent’s only substantive counter-argument was the observation, surprising to me, that the trial judge had had no difficulty discerning the phrase, and the defence had made no objection to it. To back this up, he provided the voir dire from the trial (which I had not previously seen). It is worth looking at this in a little more detail.

5.3.1 The trial judge had no difficulty discerning the phrase

The voir dire shows defence counsel objecting strenuously to admission of the audio, on grounds that parts of it are indistinct and the transcript is potentially suggestive. To demonstrate this, he asks the judge to listen to the audio without the transcript. The judge responds that there is no point in his doing that, since he is hard of hearing, and without the assistance of the transcript will hear nothing. This makes the Crown’s assertion ‘His Honour had no difficulty discerning the phrase’ seem absurd. It is important to recall, however, that this audio is so indistinct that even those with good hearing, and good equipment, hear essentially nothing in the absence of a transcript (see Fraser & Stevenson, 2014).

So the judge reviews the audio with the assistance of the transcript, unintentionally providing a textbook example of a listener being primed by a transcript:

\[
\text{HIS HONOUR: \ldots when I stopped the tape and asked for it to be played again, I was amazed how much more I heard when I suddenly focused exactly on what I was listening to.}
\]

However, in recognition of his own hearing difficulty, he consults the Crown for an opinion:

\[
\text{HIS HONOUR: As I said my ears aren't necessarily the best. Mr Crown, I assume you have listened to [this tape] a number of times?}
\]

\[
\text{CROWN PROSECUTOR: Yes, I have.}
\]

\[
\text{HH: [D]o you have normal hearing, can I inquire?}
\]
CP: Yes, I do. I did have a hearing problem two years ago but it’s resolved itself, but I do now have good hearing.

HH: Have you conducted a check yourself as to whether what appears [in the transcript], to your ears, also appears on the tape?

CP: Yes, I have … [T]here was part of that passage at the bottom of page 5 I thought I missed when I first heard it [but when I listened again I heard it quite distinctly]. Apart from that, the rest of it I heard quite distinctly.

In pursuit of fairness, the judge then asks defence counsel to use the lunch hour to highlight any sections of the transcript he finds particularly indistinct, so that he (the judge) can review them more carefully in the afternoon session. Evidently no one realised that this process of highlighting, by reinforcing the priming effect of the transcript, can be guaranteed to have the opposite result to the one intended – and indeed, after lunch defence withdrew their objection.

To linguists, perhaps the most interesting aspect of the voir dire is what it omitted. The barristers did not provide, and the judge did not require, any questions for the detective regarding his own track record with transcription. Such questions might have revealed that, some years earlier, another murder conviction had been obtained on the basis of an inaccurate transcript by this very detective. One major difference from the ‘pact’ case was that in that earlier case the unreliability of the transcript had been discovered in time for an appeal, which, by coincidence, was heard by the very judge now trying the pact case.

Another major difference was that in the earlier case, the judge accepted expert evidence that the police transcript was wrong, and quashed the conviction. However, there were no repercussions for the detective, who provided evidence in the ‘pact’ case (before the same judge), and indeed continues to this day providing transcripts to ‘assist’ juries reaching verdicts on major crimes (Fraser, 2018a). Nor has the appeal case had any effect on standard legal practice for admission of police transcripts, despite showing that it has the potential to mislead juries (see Fraser, 2018b).

5.3.2 The defence made no objection to the ‘pact’ phrase at the time

The voir dire also revealed that, despite objecting to the audio and to the transcript in general, the defence had explicitly accepted the ‘pact’ phrase itself:

DEFENCE: I’m not objecting to that portion. The part about the pact, I am not objecting to that.

This, along with the defence submissions (discussed below), helped me to understand some puzzling experiences. Throughout the application process I had been surprised at how little support I received from the defence. To be honest, I thought privately this might be because the application involved some admission of fault on their part (even though I stressed repeatedly that any fault is with the legal process, not with the individuals following it). My thoughts seemed to be confirmed
when the pace picked up after the application process was taken over by a new barrister, who had not been involved in the original trial.

However, the *voir dire* suggested another explanation, which turned out to be correct. The new barrister confirmed that the defence had accepted the ‘pact’ phrase during the trial, and indeed that the reason for the earlier go-slow with the application was that they had not accepted my analysis that it was inaccurately transcribed. He described himself listening to the audio with the defence team, all agreeing ‘sure sounds like at the start we made a pact to me’. It was only after being exposed to ‘It’s fucking payback’ that he himself came to doubt the ‘pact’ phrase – though he confessed he really accepted the alternative only after asking his children, with their keen young ears, to listen to the audio – and seeing them blush at the swearing.

Unfortunately, he now considered the new phrase to be the actual words spoken. Indeed, in submitting the application for review, his submissions described my report as providing ‘fresh evidence’ that the applicant did not say ‘at the start we made a pact’, but rather ‘at the start … fucking payback’. This is wrong in two important ways. First, I had emphasised that ‘it’s fucking payback’ was not suggested as an accurate transcript, only as a more plausible alternative (a contradiction picked up by the Crown respondent). Second, I had suggested ‘it’s fucking payback’ as an alternative for the whole phrase ‘at the start we made a pact’ (not just the last part). The way the defence submissions characterised it downplayed the implausibility of the ‘pact’ phrase that my report had emphasised (an error not picked up in either the Crown response or the Decision).

5.4. *The Decision*

The judge’s decision started with a useful statement confirming what would be needed for the application to succeed, namely:

> *an expert opinion that the transcripts of the relevant recordings were in fact wrong and that this could be convincingly demonstrated. [...] In such a case there would be a real risk of injustice to the applicant because the directions given to the jury would not have been adequate in the face of a factual error that would have meant that the aide memoire was not even arguably correct but was on the contrary entirely or at least partially inaccurate and misleading.*

To my mind this is exactly what my report did show. However, the judge did not see it this way, and rejected the application on both criteria.

5.4.1 *Not fresh evidence*

His Honour read my report as merely raising ‘a possibility that the jury were misled because one cannot be certain that the words in question were accurately recorded’ and as pointing to ‘studies that suggest that seeing an inaccurate transcript can significantly influence a listener’s perception’. Of course my reference was not to general ‘studies’ vaguely ‘suggesting’ a possibility that listeners ‘can’ be influenced by transcripts, but to specific experiments demonstrating empirically that listeners to
this particular audio actually are influenced by this particular transcript, which actually is inaccurate. Nevertheless the judge argued that potential influence by a transcript is ‘a possibility that is well recognised and implicitly anticipated by the standard jury directions’, which were properly given by the trial judge. Thus my report, though ‘interesting and compelling’, gave him no sense of disquiet.

What scholars of linguistics might find disquieting is his suggestion that he could have been persuaded the transcript was factually wrong if I had used ‘enhanced listening technology that established or suggested that what went to the jury was incorrectly transcribed’. It is not entirely clear what he means by ‘enhanced listening technology’, especially since the audio admitted at the trial was itself ‘enhanced’, albeit ineffectively. However it does give the impression it might not be only juries who suffer from the ‘CSI effect’, i.e., over-confidence in the power of ‘forensics’ as depicted in television crime shows (Holmgren & Fordham, 2010).

After some additional confident but poorly informed discussion of priming (attributing views to me that I did not state in my report, and do not hold), he concluded that my report was not fresh evidence, and thus failed on the first criterion for achieving a review of the conviction.

5.4.2 Would not have made a difference

The judge further agreed with the Crown that, even if my evidence had been presented at the original trial, it would have made no difference to the verdict – meaning the application failed also on the second criterion. This, he argued, was because the rest of the recording contained ‘other incriminating material’ supporting the prosecution case, which, by inference, would have persuaded the jury of the defendant’s guilt even if they had accepted my arguments that he did not say ‘at the start we made a pact’.

This is an especially interesting argument in that it seems to contradict not just my expert evidence that the police transcript was unreliable and misleading throughout, but the law’s own principle that the ultimate decision as to the content of indistinct audio rests with the jury. In assuming that the incriminating material in the rest of the transcript would have supported their verdict even if I had given my evidence, the judge is assuming the jury accepted the police transcript as presented.

Even more disquieting, the same thing happened in the unsuccessful High Court appeal mentioned in Section 5.1 above, where an excerpt from the police transcript containing the ‘pact’ phrase was read out verbatim, with no demur from either defence or judge, as demonstration that the defendant understood the nature of a joint criminal enterprise (an essential plank in the prosecution’s successful argument for rejecting the application to appeal).
5.5. What can we learn from the case study?

While the lessons of this case study seem generally clear, it may be worth mentioning, in closing this section, two matters that sometimes arise in discussion of this case.

First is the common question: ‘Why didn't the defendant himself dispute the ‘pact’ phrase, if he didn't say it?’ The answer is that this case powerfully illustrates the fallacies discussed in Section 4.3 above. Recall that the ‘pact’ phrase is 2.5 seconds within an (extremely indistinct) 38-minute recording. When the defendant was asked to check the police transcript against the audio, he focused (understandably enough) on identifying errors that supported his plea – that he was guilty of being an accessory after the fact, but not guilty of masterminding a joint criminal enterprise to commit murder. Indeed he noted many such errors (some phonetically valid, some not). However, from his perspective, the ‘pact’ phrase was not an issue, since it supported his own version of events: that the killer (his son) had confessed the murder to him, and he (the father), while urging the son to turn himself in, had pledged to support him no matter what.

The point of course is the flaw in the legal concept that reliable interpretation of indistinct audio can be achieved by having a transcript produced by an (untrained and inevitably non-objective) police officer, then ‘checked’ by an (untrained and inevitably non-objective) defendant, and evaluated by an (untrained and inevitably non-objective) jury. As discussed further below, covert recordings used as evidence in court need to be accompanied by transcripts produced by independent professionals following a validated process.

Second is the question: ‘Would the application for review have been accepted if an expert had simply asserted (perhaps citing ‘enhanced listening techniques’) that ‘it’s fucking payback’ as the correct transcript, rather than hedging about it being only a more plausible alternative?’ Indeed this might have allowed the application to go ahead – though it could still have failed the second test (see Section 5.4.2 above). However, it is hard to agree this would have been a better strategy. For one thing, the fact the opinion is scientifically unjustified (see Fraser 2017b) might have been picked up during a full review, and, if not, would have created a bad precedent for future trials. For another, some listeners who have (incorrectly) accepted ‘it’s fucking payback’ consider it to be even more incriminating than ‘at the start we made a pact’ – which surely supports the principle that it is better to exclude untranscribable audio than to allow plausible guesses.

6. A problem in need of a solution

Surely even one example of a manifestly inaccurate transcript being so confidently accepted by a court as ‘assistance’ in interpreting such important evidence would be evidence of a major problem in the legal system. In fact, there are many examples. The case discussed here is merely one in which a perfect storm of factors created a particularly bad outcome (made more poignant by the fact the protagonist died in...
prison shortly after the application was rejected). However, there is nothing to say these factors will not occur, or have not already occurred, in other cases.

But that is not end of the problems caused by reliance on police transcripts. The transcript of a covert recording is the crucial foundation for everything that is done with the evidence contained in the audio (Fraser, 2018b). For example, readers who engage in forensic case work should note that police transcripts are sometimes sent to linguists in hopes expert opinion can back up a particular interpretation of what speakers mean by the words transcribed. It is highly advisable, before complying with such a request, to check the provenance, and reliability, of any text being analysed.

As another example, police transcripts typically provide 'speaker attribution', i.e. indication of which participant produced each utterance, and police give speaker identification evidence as 'ad hoc experts' – typically as part of the ‘Statement of Facts' in the brief. Again, speaker attribution and identification are specialist tasks which, when undertaken by amateurs, are highly prone to error (French & Stevens, 2013; Rose 2005). Even if an expert is subsequently consulted, the police attribution typically forms the basis of the expert's analysis. Without extremely careful handling, this creates substantial risk of the ‘confirmation bias’ to which even the most objective of forensic measurements are prone (cf. Thompson, 2009).

Given the strong reliance inevitably placed on the transcript, it is essential that all indistinct audio used as evidence should be accompanied by a reliable transcript. The question now is: how to achieve this.

7. What not to do

One solution that readily comes to the minds of linguists is to educate police and/or lawyers in how to create and use transcripts in a more effective manner. Unfortunately this is not likely to solve the problems effectively. One reason is that training of this nature can serve to increase practitioners' confidence in their own ability, without necessarily improving their actual ability (cf. D. Harris, 2012, on training in lie detection). So, for example, defence barristers have responded to education about priming by committing to check police transcripts extra carefully, while judges have decided that indistinct audio should be played without any transcript (this of course leaves open the strong possibility of priming by context and suggestion (cf. Fraser & Stevenson, 2014)).

Another solution commonly suggested is that all forensic audio should be transcribed by experts in linguistics. This is not my recommendation (see Fraser 2015; 2018b). For one thing, it creates a number of practical problems, including prohibitive expense, insufficient numbers of genuine experts willing to take on the work, and difficulty in ensuring the courts distinguish appropriately between genuine experts and those unjustifiably claiming ‘specialised knowledge’ (cf. Edmond, 2015).
There are also more linguistically interesting reasons to reject this suggestion. Though many branches of linguistics use transcripts of various kinds as the basis of analysis, these are not suitable to be applied directly in the context of forensic transcription. Far better is to combine deep understanding of general principles affecting the reliability of transcripts under various circumstances (which may require development of a more general theory of transcription than is currently available) with deep understanding of the context in which forensic transcripts are used (see extended discussion in Fraser, 2014).

8. What's needed

Ultimately, the solution is to handle forensic transcription like analysis of other scientific evidence, via a process that does not use police interpretation as its starting point, and does not rely on direct evaluation of data by lawyers and juries (see further discussion in Fraser, 2018b).

Achieving this requires a collaborative effort. While it is clearly essential for law and law enforcement personnel to gain a better understanding of linguistics as a science, it is also necessary for linguists to increase their understanding of the legal context – not just about how it is supposed to work but about how it really does work. This requires linguists to do more than just respond to requests for ‘specialised knowledge’ as they happen to come in. To reveal poor practice, and understand its rationale, it is necessary to enquire as to the basis for the request – and even more importantly, to investigate what happens in those many cases where linguists are not asked for their expertise.

As an example, my own early work on forensic transcription, accepting the doctrine that the transcript was for the assistance of the jury, wasted time demonstrating that juries can be misled by inaccurate transcripts (Fraser et al, 2011). This of course missed the mark, since lawyers are well aware of that fact. It took digging beneath the surface to realise that the real issue is not juries’ perception but lawyers’ perception (as discussed in Section 4.4 above).

9. Conclusion

Due to the strength of legal authority, the poor practice discussed above (along with related practices outlined in Fraser 2018b) will not change in response to external advice, no matter how expert or how well supported by evidence. It is really necessary to create a will to change within the criminal justice system itself, and to build recognition that developing an effective solution requires collaboration with linguists. Hopefully the present paper has created unease and disquiet among linguists sufficient to motivate participation in a concerted effort, first to explain why the current situation is not acceptable, and then to contribute to development of effective solutions. The recent Call to Action delivered to the Australian Institute of Judicial Administration by the Australian Linguistics Society and the Applied
Linguistics Association of Australia makes a promising start, but there remains much work to do.

Beyond the practical imperative, forensic transcription raises interesting theoretical questions which can be missed if forensic transcription is treated as a mere ‘application’ of existing knowledge. Two examples have been alluded to which may have relevance to various branches of linguistics. One is the need for a broader theory of the nature of transcription, and the factors that affect the reliability and usefulness of transcripts under various circumstances. Another is the need for a deeper theoretical understanding of the false beliefs of educated common knowledge, and development of evidence that can help to overcome them.

10. Acknowledgments

The present paper has benefitted indirectly from discussion with, and directly from comments by, a number of people, to all of whom I owe thanks: Alex Bowen, Kate Burridge, Felicity Cox, Diana Eades, Nick Enfield, Nick Evans, Peter French, Rod Gardner, Peter Gray, Georgina Heydon, David Nash, Phil Rose, Cindy Schneider, Roger Shuy, Lesley Stirling, Nick Thieberger, participants in several seminars, and two anonymous reviewers. Of course, I remain solely responsible for any inaccuracies or infelicities.

11. References


Fraser, H. (2018b). Thirty years is long enough: It’s time to create a process that ensures covert recordings used as evidence in court are interpreted reliably and fairly. *Journal of Judicial Administration*.


