

*Please note: This is a summary of the main article. The text is substantially different. If you wish to cite the findings, it is better to refer to the main article in IJSLL.*

## **Why evaluation of forensic transcripts should not be left to the jury**

**Helen Fraser**

<http://helenfraser.com.au/forensic>

---

### **Abstract**

A forensic transcript is a representation, in writing, of speech captured via a covert recording (telephone intercept, listening device, etc.) and used as evidence in a legal case. When the clarity of such a covert recording is poor, there can be dispute between prosecution and defence over the appropriate transcription of particular words or sentences (in one famous example, there was dispute as to whether a defendant had said 'I shot a man to kill' or 'I showed a man a ticket').

The present paper is a summary of a longer article (Fraser, H., Stevenson, B., & Marks, A. to appear. Interpretation of a crisis call: The persistence of primed perception of a disputed utterance in a forensic transcription. *International Journal of Speech Language and the Law*, 18.2), which reports an experiment whose results provide graphic evidence of why resolution of such disputes should not be left to the jury.

Note: A short (5-10 minutes) version of the experiment, collecting no data or personal details, is available via <http://helenfraser.com.au/forensic>. Readers are encouraged to follow this link to form their own opinion about the materials before reading further.

### **Introduction**

Legal cases increasingly include evidence in the form of audio recordings obtained from telephone intercepts, covert listening devices and similar sources (Coulthard & Johnson 2007). Sometimes these recordings are of poor quality, to the extent prosecution and defence disagree as to their content, either throughout, or in particular 'disputed utterances'. In such situations, debate can surround not just the accurate interpretation of what words (if any) are spoken in the recording, but the admissibility of the recording as evidence for presentation to a jury.

The case of David Bain, from New Zealand, offers a good example (see [http://en.wikipedia.org/wiki/David\\_Bain](http://en.wikipedia.org/wiki/David_Bain) for further information). In this case, a very brief and barely audible section of a recorded crisis call was alleged by prosecution to be the caller confessing to murder, and by defence to be uninterpretable speech, perhaps not even speech at all.

How can the truth be decided in such a situation? One possibility is to play the recording to the jury, perhaps with advice from expert witnesses on each side, and let them decide. This approach raises a problem, very well known to the disciplines of phonetics and psycholinguistics (Byrd & Mintz 2010) and acknowledged to a certain extent by the legal system: the danger of ‘priming’ (i.e. the tendency, especially with a degraded signal, for the ear to hear words that have been suggested). This means the very act of telling the jury one of the interpretations to be evaluated might cause them to hear that particular interpretation. Further, once they have heard that interpretation ‘with their own ears’, it might be difficult for them to ‘unhear’ it sufficiently to consider an alternative interpretation with equal objectivity. Indeed, the primed perception might subconsciously affect their opinion of the defendant and the case as a whole, even if it were later shown to be an incorrect interpretation of the disputed utterance.

Based on arguments like these from the defence, the disputed utterance in the Bain case was not put to the jury, so we will never know for sure what would have happened. However, the present paper offers a summary of findings from an experiment designed to track participants’ interpretation of the actual disputed utterance from the case, as it is progressively affected by evidence that might have been produced for the jury. The results show the effects of priming may be considerably greater than is usually recognised.

## **Background**

For relevant background, it is hard to improve on the account provided by (Innes 2011).

On 20 June, 1994, in Dunedin, New Zealand, a young man returned from his early morning paper run to find a shocking sight: his family had been murdered. David Bain was the only survivor, his parents, two sisters and younger brother having been shot during the time David was out delivering newspapers in his local area. Some 30 minutes after returning home, the distressed 22 year old rang the emergency services for help. Three days later, he was arrested and charged with the murders. After a three week trial before a jury in the Christchurch High Court in 1995, he was duly convicted and sentenced to life imprisonment with a non-parole period of 16 years.

[...] Bain appealed his conviction. In the event, it was considered three times by the Court of Appeal. [...] As a final step, the defense team applied for leave to appeal to the Judicial Committee of the Privy Council in London, which was duly granted. The Privy Council decision, issued in May 2007, some thirteen years after the original conviction, [...] quashed the convictions and ordered a retrial [...]. Bain was released from prison. The Solicitor General in New Zealand, after some time (and much community discussion as to whether he should), duly ordered a retrial.

[...]

Shortly after the Privy Council decision [in 2007], the police took the recording [of Bain's original call to emergency services] to a sound studio to be digitised. [...] A detective then heard words on the recording which had never been heard before, and therefore had not been part of the evidence in the first trial. These words were to become a bone of contention because, if they had indeed been spoken, they would have been taken as an admission of guilt and thus would be crucial in the upcoming trial. The next step taken (October 2007) was to ask the ambulance officer who had answered the call whether he thought the words had been said. On listening to the digitised recording, he said that he could hear the words and that he was 'stunned' that he had not heard them before [...].

(Innes 2011)

Advice was sought, by both prosecution and defence, from several speech experts (again, see Innes 2011 for details). None of these experts found any evidence to support the 'I shot the prick' interpretation, and, though they provided somewhat differing opinions as to what (if anything) might have been said in the disputed section of the call, all were united in emphasising that it could not be transcribed with

confidence, and recommending caution in presenting any interpretation to a jury, due to the danger of priming.

This led to a great deal of legal argumentation, through a further series of appeals, about the admissibility of the disputed utterance as evidence. The final appeal was heard by the New Zealand Supreme Court on 2 March 2009.

Counsel for Bain argued that the sounds concerned were irrelevant as they did not prove anything and would create unfair prejudice. She was concerned that judicial direction of the jury would not be enough to avoid the potential dangers. She further contended that if the disputed portion did not contain words, then it was not probative and therefore was irrelevant on that basis also. She also argued that the expert evidence should not be included because it carried the risk of suggestibility in itself; it could have had the effect of priming the jury.

Counsel for the Crown argued, on the other hand, that the accused's intent was crucial here, given that the experts could not resolve the matter, and that the surrounding context was relevant in interpreting the recording. In fact, he argued, the existence of competing interpretations was not grounds for withholding the evidence from the jury. The Crown firmly believed that the disputed portion of the recording contained an admission and that there was no question but that this was relevant (as neither authenticity nor identity were in dispute).

(Innes 2011)

Following intricate legal argument, the Supreme Court allowed Bain's appeal, but ruled that, when the retrial was held, the disputed section should not be played to the jury.

The ambulance call was duly played with the disputed portion excised. The call was stated to be relevant as to the state of mind of the young man who would be charged with the murders some three days after the murders. The jury's verdict? Not guilty. After 13 years in prison, David Bain was free to go.

The final legal act in the matter of the telephone call was an application by the media, immediately after the trial concluded, for the suppression order to be removed. [...] Bain, unsurprisingly, applied to have the suppression remain. [...] The Supreme

Court did not agree that the private interest of Bain outweighed the public interest [...].

This led, of course, to headlines ‘I shot the prick’, and feverish media and public interest. Anecdotally speaking, the general public appeared to be divided on hearing the words alleged by the police: some could hear the words, others could not. (Innes 2011)

Regardless of opinion as to the justice of lifting the suppression order, it has had one good outcome: the audio recording is now in the public domain, allowing us to explore the question of just what a jury might have made of the material if it had been presented in court.

### **The experiment**

The experiment summarised here used the actual crisis call from the Bain case, and took the form of an online survey, in which participants listened to the call and the section of interest, then were invited to imagine they were on the jury as ‘evidence’ was presented in several stages. (Of course this evidence did not replicate the facts of the Bain case in detail, as the intention was to design a statistically valid experiment with general implications, not to make statements about the Bain case in particular.) At each ‘Evidence Point’, participants were asked to state what they thought was said in the section of interest, and provide their level of confidence.

The survey was taken by 190 people from Australia, New Zealand, the UK and the US. At the third Evidence Point, participants were randomly divided into two groups, each receiving parallel but slightly different evidence. Specifically, Group A was given a story in which suspicion was cast on the caller, and the section of interest was alleged to contain the words ‘I shot the prick’, while Group B was given a story that cast suspicion on the caller’s father, and alleged the section of interest contained the words ‘He shot them all’. This second phrase was invented as a foil, intended to be semantically similar to the actually alleged phrase, but a poorer match to the acoustics of the section of interest.

At the end, participants from both groups were given the full story, and asked to provide a Final Verdict, by rating interpretations for the section of interest on a 5-point scale likelihood. They were also asked to indicate, again on 5-point scales, their conclusion as to the guilt of each of: the caller, the father, someone else.

**Table 1**

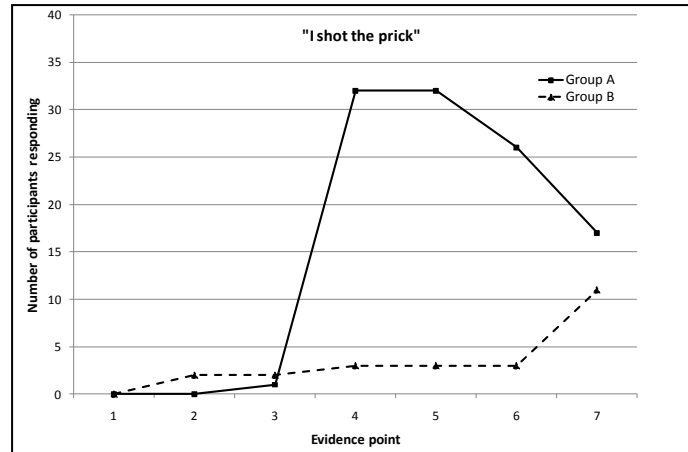
Outline of survey structure

	Group 1	Group 2
Initial questions	Presentation of the crisis call as a whole, with questions about intelligibility of call and trustworthiness of caller	
Evidence 1/Baseline	Introduction to the section of interest, and indication it is part of a genuine crisis call	
Evidence 2	The crisis call is part of a murder case and the section of interest is relevant	
Evidence 3	The section of interest contains an admission of guilt by the caller	The section of interest contains an admission the caller knows his father is guilty
Evidence 4	The section of interest is alleged to contain specific words 'I shot the prick'	The section of interest is alleged to contain specific words 'He shot them all'
Evidence 5	An audio engineer backs up the allegation the words are 'I shot the prick'	An audio engineer backs up the allegation the words are 'He shot them all'
Evidence 6	A phonetics expert refutes the audio engineer's evidence	A phonetics expert refutes the audio engineer's evidence
Evidence 7/Final Verdict	Both groups receive the full story, including general expert opinion that the section of interest does not contain the words 'I shot the prick' and the caller was exonerated of the crime. Both groups rate interpretations of the section of interest, and provide Guilt ratings for all suspects.	

## Results

**Figure 1**

Results for 'I shot the prick' (see the full paper for results for other interpretations)



The most important findings are shown in Figure 1. Until Evidence 3, the groups are statistically similar, with 'I shot the prick' heard by very few participants in either group (see below for further discussion of these few individuals). However, when Group A were primed with 'I shot the prick' at Evidence 4, about one-third of them reported hearing it. This number remained statistically similar through the next two Evidence points (audio engineer and phonetics expert evidence), and dropped after the full story. However, despite the strong indication given by the full story that the 'I shot the prick' interpretation was incorrect, 17 participants (18%) still preferred 'I shot the prick' at Final verdict.

In Group B, very few participants heard the phrase 'I shot the prick' even after Evidence 4, as this group was primed with a different phrase, so the groups are statistically different during Evidence 3-6. However, at Evidence 7, they return to being statistically similar. This is because 11 Group B participants (12%) offered 'I shot the prick' as their interpretation of the disputed utterance at Final Verdict – clearly influenced by mention of the phrase (for the first time for this Group) in the full story. Again, this was despite the fact that this mention was in the context of an

explanation that, though this interpretation had been alleged, it had later been shown to be unreliable, and that the caller had been found not guilty.

As part of the Final Verdict question, participants were asked who they thought was guilty (caller, father or someone else), with each suspect to be rated on a scale where 1 = guilty, 3 = not sure/can't tell, 5 = not guilty. Overall, all three suspects received a similar mean guilt rating of around 3.65 (mode = 4), with no statistical difference between groups or suspects. Importantly, however, those participants (from both groups) who responded 'I shot the prick' at Evidence 7 were significantly more likely than others to find the caller guilty.

Another important finding was the strong correlation, across both groups, between participants' initial lack of trust in the caller (measured at the very beginning of the experiment, when the crisis call had been heard but no information about its context had been given), and their final attribution of guilt to him. In other words, participants who did not like the caller's tone of voice at the outset were more likely to consider him guilty, despite having been told he had been exonerated by the court).

### **Demographic analyses**

Correlations of results with demographic categories were run. In general, participants behaved similarly, regardless of gender, age, educational background or dialect. However, there are a few effects worth mentioning.

#### Phonetics background

In both groups, professional phoneticians gave zero 'I shot the prick' responses throughout all Evidence points, and gave a significantly higher percentage of 'Uninterpretable speech' responses at Final Verdict.

#### Prior knowledge of the Bain case

Only four participants heard the phrase 'I shot the prick' without it having been suggested to them (i.e. before Evidence 4 for Group A or Evidence 6 for Group B). Of these, one later indicated prior knowledge of the case via an undergraduate class in forensic science, while two (50%) belonged to a particular subgroup, discussed next.

### Experience in a court case

Six respondents indicated they were police officers (all Australian, and all stating they had no prior knowledge of the Bain case). Of these, two heard 'I shot the prick' before it was suggested. Further, three (50%) gave a Final Verdict interpretation of 'I shot the prick' – a statistically significant difference from the 14% of participants overall. Finally, these six police officers gave a higher average 'caller guilt' score (2.83 compared to 3.65 for all participants), with two (33%, compared to fewer than 8% of participants overall) finding the caller to be 'definitely guilty' at Final Verdict (i.e. after being told he had been exonerated). The numbers of police are too small to enable robust conclusions to be drawn, but a little more will be said on these observations below.

### **Discussion**

These results demonstrate why phonetics experts express concern about the practice of leaving evaluation of forensic transcripts of poor quality recordings to a jury. The experimental evidence provided here indicates that, had the disputed utterance in the Bain case been put to the jury, around a third of jurors might have 'heard' the phrase 'I shot the prick' when it was suggested to them – with significant implications for their likelihood of finding the defendant guilty – even though there is no phonetic evidence to support this interpretation, and virtually no-one has ever heard it un-primed. Further, though there is not space to go into it here, it is worth noting that the detailed analysis in the full article provides evidence of more subtle priming effects which might have influenced the interpretation of even those jurors who did not accept the primed interpretation in full.

The evidence also indicates that more than half of the 'primed' jurors would have persisted in their interpretation even if expert evidence had been provided to show that it was incorrect. This suggests that a mere caution from the judge would have been unlikely to return these jurors to a neutral frame of mind suitable for evaluating alternative interpretations, had these been provided.

None of this means that disputed recordings must never be presented to a jury. It does suggest that, wherever possible, potential disputes over the transcription of

covert recordings should be discovered and resolved at the voir dire stage, before they are entered as evidence in a legal case (see Fraser 2010, or the summary in Fraser 2011). In most cases, this can be achieved by ensuring that all potentially controversial transcripts are properly evaluated by a genuine expert in cognitive phonetics. In the rare cases where disagreement remains even after such evaluation, of course the competing hypotheses must be considered by the court. However, this should be done with advice and assistance from the experts, who can ensure the material is presented to the jury in a valid manner.

The results of this experiment also indicate the reason for concern among phonetics experts about the practice of allowing police to evaluate their own transcripts as ‘ad hoc experts’. The police who took part in the present experiment were not equivalent to ‘ad hoc experts’ (having no prior familiarity with the voice of the caller), and, as already mentioned, their numbers were too low to allow firm conclusions to be drawn. Nevertheless, their apparent willingness (even greater than that of the experimental population as a whole) to accept the incriminating interpretation of the disputed utterance in the face of expert evidence that it was incorrect, and base a ‘verdict’ of guilty on their interpretation, surely gives pause. This is certainly not intended a criticism of police themselves. The effects of priming are generally subconscious. If the results of this experiment are upheld by further investigation, the cause would likely be, not intentional bias, but a system which, by giving police powers to present evidence as ‘ad hoc experts’, may encourage them to place too much credence in their own untutored perception. The solution then would be as simple as ensuring greater education of police about the fallibility of speech perception (see Fraser 2011).

## **Conclusion**

In the Bain case, as in others, the responsible conclusion is that the disputed utterance is ‘untranscribable’. This does not mean it is not possible to have an opinion as to what might have been said. Rather, it means that no opinion can be validated to a level of certainty appropriate to a high stakes context such as a legal case. This may be frustrating to the legal system, but is really no different to the situation that obtains with other forensic sciences. For example, if experts consider a

smudged fingerprint too degraded in quality to provide reliable evidence, we do not encourage juries or 'ad hoc experts' to form their own opinion whose print it might be. Rather we ensure that the smudged print is not entered as evidence. The difference is that, with speech, non-experts have such a strong and confident (though often unreliable) sense of particular words being 'there to be heard' that it can be hard for them to accept expert evidence to the contrary.

The experiment reported here provides strong support for the conclusion finally reached, on appeal, by the New Zealand Supreme Court in 2009, that the disputed section should not be presented to the jury. This is in addition to the anecdotal support already given (cf. Innes 2011) by the fact that release of the recording to the media created a frenzy of uninformed public vitriol after the trial had ended.

What is urgently needed now is widespread adoption of a standard practice whereby the legal system can reach similarly responsible conclusions in similar cases – with far lower costs in terms of time, money and personal suffering than in Bain and other cases (cf. Fraser 2010). Hopefully this can be achieved in the near future through close cooperation between the legal system and experts in cognitive phonetics (cf. Gray 2010).

## **References**

- Byrd, D and Mintz, TH. 2010. *Discovering Speech, Words and Mind*: Wiley-Blackwell Publishing.
- Coulthard, RM and Johnson, A. 2007. *An Introduction to Forensic Linguistics: Language in Evidence*. London and New York: Routledge.
- Edmond, G. 2010. Impartiality, efficiency or reliability? A critical response to expert evidence law and procedure in Australia. *Australian Journal of Forensic Sciences*:1-17.
- 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* 10:203-226.

- Fraser, H. 2009. The role of 'educated native speakers' in providing language analysis for the determination of the origin of asylum seekers. *International Journal of Speech Language and the Law* 16:113-138.
- Fraser, H. 2010. Transcripts in the legal system. In *Expert Evidence (Chapter 100)*, ed. I Freckelton and H Selby. Sydney: Thomson Reuters.
- Fraser, H. (2011). Forensic Transcription. Presentation at Expert Evidence Conference, National Judicial College of Australia and ANU College of Law, Canberra, 12-13 Feb 2011. (<http://helenfraser.com.au/forensic>)
- Fraser, H., Stevenson, B., & Marks, A. (to appear). Interpretation of a crisis call: The persistence of primed perception of speech. *International Journal of Speech Language and the Law*, 18
- Frith, C. 2008. *Making up the Mind: How the brain creates the mental world* Oxford: Blackwell.
- Gray, PRA. 2010. The expert witness problem. *International Journal of Speech Language and the Law* 17:201-209.
- Innes, B. 2011. R v David Bain: A unique case in New Zealand legal and linguistic history. *International Journal of Speech Language and the Law*. 18: