



'Assisting' listeners to hear words that aren't there: dangers in using police transcripts of indistinct covert recordings

Helen Fraser

To cite this article: Helen Fraser (2018) 'Assisting' listeners to hear words that aren't there: dangers in using police transcripts of indistinct covert recordings, Australian Journal of Forensic Sciences, 50:2, 129-139, DOI: [10.1080/00450618.2017.1340522](https://doi.org/10.1080/00450618.2017.1340522)

To link to this article: <https://doi.org/10.1080/00450618.2017.1340522>



Published online: 06 Jul 2017.



Submit your article to this journal [↗](#)



Article views: 118



View related articles [↗](#)



View Crossmark data [↗](#)



'Assisting' listeners to hear words that aren't there: dangers in using police transcripts of indistinct covert recordings

Helen Fraser

Independent scholar

ABSTRACT

Results are reported of a new experiment using an indistinct covert recording from a real murder trial, along with the police transcript admitted to 'assist' the court to hear its contents. Previous research using the same material has shown that the police transcript is inaccurate, yet nevertheless highly influential on the perception of listeners 'primed' by seeing words it suggests. The current experiment examines the effects of priming participants with a made-up phrase that vaguely fits the acoustics of one section of the recording. Results indicate that a very high proportion of listeners are easily 'assisted' to 'hear' the made-up phrase. Discussion argues that audio of this quality should only be used as evidence if accompanied by a reliable independent transcript.

ARTICLE HISTORY

Received 5 June 2017

Accepted 7 June 2017

KEYWORDS

Forensic; phonetics; transcription; covert recordings

Introduction

Ever-increasing numbers of criminal trials feature evidence in the form of lawfully obtained covert recordings (conversations captured, most often on behalf of police, without the knowledge of one or more participants). Due to the difficulty of controlling covert recording conditions, the audio is often of extremely poor quality. Current Australian law allows police (in the role of so-called 'ad hoc expert') to provide transcripts to 'assist' the jury in making out what is said (for background, see *Real forensic experts should pay more attention to the dangers posed by 'ad hoc experts'*, Guest Editorial, this issue). The current paper reports a short experiment building on previous research demonstrating the dangers of this practice.

Aim

This very simple experiment uses audio from a real murder trial, along with a phrase from the police transcript that played a crucial role in obtaining a guilty verdict. Previous research¹, unfortunately conducted too late to affect the verdict, has shown, first, that the detective's transcription of the phrase is inaccurate, and second, that the phrase nevertheless has a strong priming effect on listeners. This means that, especially in the context of background knowledge or assumptions about the case, it causes listeners to 'hear' the inaccurately transcribed phrase and use it in forming opinions about the speaker's guilt.

The current experiment uses the same audio, this time aiming to demonstrate how easy it is for a transcript to persuade listeners they are ‘hearing with their own ears’ words that have simply been made up for the sake of an experiment. This is relevant to the case in which the audio was used, but also more generally, since the quality of both the audio and the transcript are typical of those admitted as evidence in Australian criminal trials on a weekly basis.

Method

Materials

The experiment uses a one-minute excerpt from an indistinct covert recording that featured as evidence in the previously-mentioned murder trial. This one minute encompasses the 14 s excerpt used in previous experiments. The audio is available at forensict transcription.com.au/audio for readers who wish to experience it for themselves before reading on (recommended).

In addition, two target phrases are used. The first is *Adelaide bank account*. This was chosen by the experimenter on the basis that a collection of syllables in Section 2 of the audio vaguely suggests that phrase to some listeners. The second is *At the start we made a pact*. This is the inaccurately transcribed phrase that played a crucial role in the murder trial mentioned above, and has been extensively studied in previous experiments. It is used here as a known baseline against which to compare results for the first phrase.

Procedure

The audio was incorporated into a survey prepared with Qualtrics (qualtrics.com), and deployed over the internet via various email lists and personal connections of the author.

Participants first listened to the audio ‘cold’, i.e. with no contextual knowledge beyond the suggestion it might be ‘forensic’. They were invited to listen as often as they wished, and transcribe what they heard into four boxes (labelled 00–15 s; 16–30 s; 31–45 s; 46–60 s) representing 4×15 s sections of the audio. Use of the boxes enabled some localization of the words transcribed, without requiring participants to type their own time stamps.

Participants were then randomized (by the survey software) into two groups. Each group was invited to listen again to see if they could locate a given phrase within the one minute recording. The first group (the ADL group) were given *Adelaide bank account* (the ADL phrase). The second group (the PACT group) were given *At the start we made a pact* (the PACT phrase).

Next, they were given the opportunity, optionally, to have another go at transcribing into 4×15 s boxes.

Finally some basic non-identifying demographic information was collected: sex, age, language background, hearing ability, occupation, etc.

Participants

Due to the recruitment process, participants came from a wide range of demographic groups. The current analysis excludes results from anyone who reported having poor hearing, being

a non-native speaker of English (no matter how proficient) or having prior experience of the audio.

This left a total of 76 participants, all native English speakers with average-to-excellent hearing, spanning both sexes, ages from teens to 70s, and a wide range of occupations, including linguists and scientists. Of the 76, 37 formed the ADL group and 39 the PACT group.

Analysis

Results were transferred from Qualtrics to an Excel spreadsheet. Responses to the Location task were simply counted. Responses to the Transcription tasks were coded by colouring relevant words and phrases (using Excel's 'conditional formatting' tool) and counting occurrences in each 15 s section.

Results

Cold transcription

This section reports how participants transcribed the audio before they had been given any suggestions as to what they might hear (cold condition).

General

Transcripts under cold condition were highly variable throughout, reflecting the extremely poor quality of the audio. The only notable commonality was that almost all participants (93%) included the word *money* in Section 3. Beyond this, however, transcripts differed wildly in both quantity and content (see Table 1 for a sample).

Table 1. A selection of participants' transcripts of Section 3 under cold condition. Note this selection excludes those that gave nothing (2), gave words other than 'money' (3), or gave just the word 'money' alone (8), then selects examples representing the range of interpretations.

be able to get money, buy me the fucking shit/shed
have any money
'now...','money';'...this year...'
now according to, we hope to get money, alert, shit
he's not going to get money, contributed, position
do you have like the moneydark and shady
we have to get money
money barbeque garden shed
now we're going to get the money...now that I'm in the fuckin' shit...and that's why I'm fucked...
Sniff. In order to get money XXX know what I mean? Blimey shit
bank account money
now according to him, he owes me money so i can cross over the income, if i can manage to find the f***** cheque
'then it's quite hard to get money, you see what I did there,'went to an emergency barbecue, it was quite shit'
now what ... you have to get money ... i can pass on that danny even ... but i managed to find it about can share
...cuz if I told them... then I couldn't.... yeah... It'd be hard to get money...
Now at work. To be holder get money...doctor croissant... Logogen doctor gee... logga chishsh demain (French word)
money, market share
now money cheated
money, barbecue package
money mortgages?
Now according to this, he helped me get money, barbeque, Amish.
in a court of law help me get money some sort of pay accounts barbecue position, definitely not

Target phrases

Here we consider the extent to which the target phrases (ADL and PACT) were transcribed under cold condition (i.e. before they were explicitly suggested).

In Section 2, one participant (of 76 = 1%) transcribed the full ADL phrase (*Adelaide bank account*), while a further 19 (25%) transcribed the words *bank account*, without *Adelaide*.

In addition, two participants (3%) transcribed *bank account* in Section 3, and two participants (3%) transcribed *bank* in Section 1.

No one transcribed anything remotely like the PACT phrase or the word *pact*, in Section 1 or anywhere else. However, in Section 1, 13 of 76 participants (17%) gave the word *back*, most often as part of a phrase such as *lying back*, *paying back*, *play back*, etc.

Location task

The ADL group

In answer to the invitation to locate the phrase *Adelaide bank account*, 25 of the 37 participants in the ADL group (68%) said they *definitely* heard the full phrase in Section 2, while a further 12 (32%) said they *thought* they heard the full phrase in Section 2.

Note, this gives a total of 100% of the BANK group affirming the existence of this made-up phrase. In addition, a further two (5%) said they also *thought* they heard the full ADL phrase in another section.

The PACT group

In answer to the invitation to locate the phrase *At the start we made a pact*, six of the 39 participants in the PACT group (15%) said they *definitely* heard the full PACT phrase in Section 1, while a further five (13%) said they *thought* they heard the full PACT phrase in Section 1.

In addition, one participant (3%) said they *thought* they heard the full PACT phrase in Section 3, while a further one participant (3%) said they *thought* they heard the full PACT phrase in Section 4. See Figure 1.

Transcription task

This section reports results of the transcription task undertaken after the localisation task (i.e. when participants had been primed by the suggestion of the phrase their group was asked to locate).

Recall that this (second) transcription task was optional (by contrast with the cold transcription, which had to be done for a participant to be included in this report).

Excluding results of participants who did not attempt to transcribe any part of the audio in this task leaves 30/37 (81%) of the ADL group and 31/39 (79%) of the PACT group. Note that these are very similar proportions. For better comparison with the cold condition, this section reports results as percentages of those who attempted the transcription ('transcribers'), rather than of the whole group.

Overall, transcripts in this condition remained highly variable in quantity and content. In both groups many participants still gave *money* in Section 3, though somewhat fewer than the 93% in cold condition (here: 83% of ADL transcribers and 90% of PACT transcribers), but again transcripts of the rest of Section 3 were highly variable (and often different from the transcripts given in the cold condition).

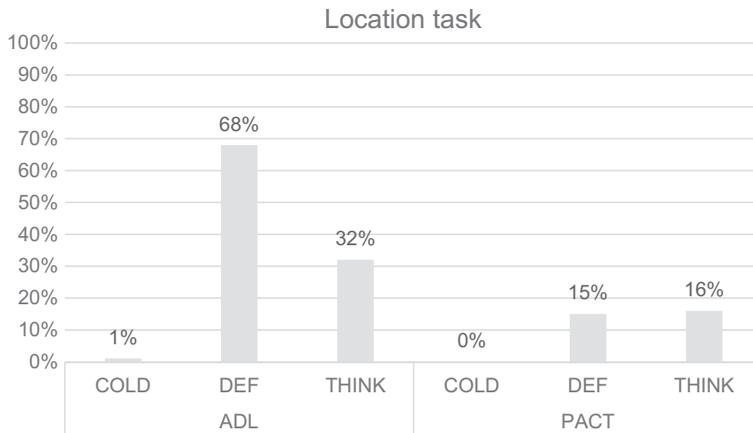


Figure 1. The number of participants who said they ‘definitely heard’ (DEF) or ‘thought they heard’ (THINK) *Adelaide bank account*’ (ADL) in Section 2, and *At the start we made a pact* (PACT) in Section 1, compared with the number who transcribed these phrases in cold condition (COLD). Note: a few participants also said they thought they heard the phrases in other sections – see text for details.

Next we consider transcription just of the phrases with which the groups were primed (See Figures 2 and 3). To avoid misunderstanding, it might be worth a reminder that the audio supporting the ADL and PACT phrases is in different sections – these are not alternative interpretations of a single utterance.

The ADL group

In Section 2, 24 of 30 transcribers primed with *Adelaide bank account* (80%) gave the full ADL phrase, while a further two (7%) gave *bank account* (none gave just *bank*).

No one gave the full ADL phrase in any other section. However, in Section 1, 11 of 30 transcribers (37%) gave the word *bank*, while five (17%) gave *back*.

No one transcribed anything remotely like the PACT phrase or the word *pact* anywhere (recall that this group had not been primed with the PACT phrase).

The PACT group

In Section 1, seven of 31 transcribers primed with *At the start we made a pact* (23%) gave the full PACT phrase, while a further two (6%) gave just the word *pact*.

One transcriber (3%) gave the full PACT phrase in Section 4.

In Section 2, four of 31 transcribers (13%) gave the full ADL phrase (recall that this group had not been primed with the ADL phrase), while a further six (19%) gave *bank account* (none gave just *bank*).

One transcriber (3%) gave the word *bank* in Section 1, while three of 31 (10%) gave *back*.

Discussion

Relation of experimental conditions to conditions in a trial

It is important to start by recognizing that participants in this experiment hear the audio under conditions far more conducive to scepticism about a suggested transcript than do a

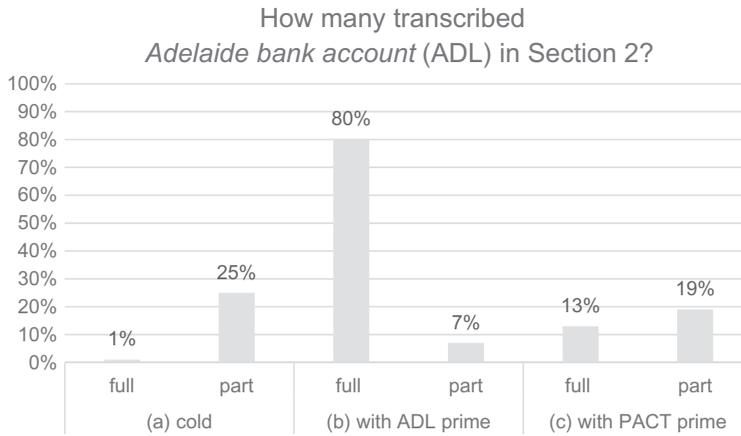


Figure 2. How many participants transcribed the full ADL phrase ('full') or just *bank account* ('part') (a) in cold condition; (b) after seeing the ADL prime (ADL group); (c) after seeing the PACT prime (= PACT group – see next section). See text for additional findings.

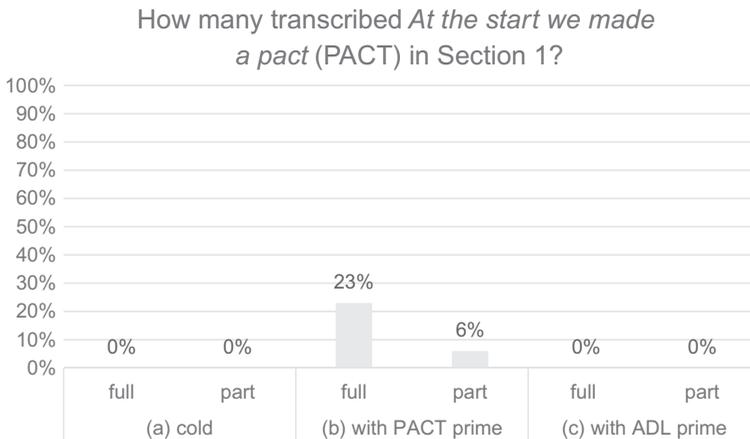


Figure 3. The number of participants who transcribed the full PACT phrase (full) or the word 'pact' (part) (a) under cold condition (both groups); (b) after seeing the PACT prime (PACT group); (c) after seeing the ADL prime (= ADL group – see previous section). See text for additional findings.

jury in a trial. First, they are selected by the recruitment process to be a relatively educated group with a certain amount of personal interest in listening and perception. Second, they are self-selected, such that those who found the task so difficult or unpleasant that they didn't even attempt the first transcription are excluded (there is no way to know how many there were). Third, they are selected by the experimenter to exclude those with less good hearing, or less proficient English.

Most important, they have been given no information about the context or content of the audio, beyond the impression that it might have something to do with 'forensics'. A jury knows a great deal of contextual background, including the purported content, long before hearing the audio (a factor known to massively affect interpretation – see the references).

Finally, participants have been given a small section of the audio, designed into an experiment that allows them to play it in full or in part, as often as they wish, under a range of conditions. These conditions include, crucially, the cold condition. This gives them the experience of how uninterpretable the audio is on its own, something a jury never gets.

The priming effect of the transcripts

Even in these relatively favourable circumstances, it is clear that both groups were heavily influenced by their priming phrase, despite the fact that both phrases are inaccurate transcripts. Although virtually no one heard either phrase in full under the cold condition, after priming, 100% of the ADL group and 31% of the PACT group located their phrase in the relevant section.

In addition, for both phrases, several participants, after priming, felt they heard the full phrase in sections *other* than the one that ostensibly supported it. Listeners do not need much acoustic evidence to confirm they 'hear' a phrase they are 'listening out for' in indistinct audio.

It is worth noting that in both groups, the number of participants who felt they could locate the suggested phrase was greater than the number who went on to transcribe it themselves (100% vs 80%; 31% vs 23%). This suggests that actually sitting down and working with the audio induces more scepticism about these transcripts than just 'listening out' for a particular phrase – although still not enough, as shown by the high proportion of transcribers who did indeed give the full phrases.

Comparative priming effect of each phrase

Both phrases had a strong priming effect, in the sense that they caused listeners who had not transcribed those words in open condition to confidently locate them in the primed condition.

However, it is clear that the ADL phrase had a far more powerful effect than the PACT phrase. The percentage of participants claiming to hear the ADL phrase after being primed with it was far higher than the percentage claiming to hear the PACT phrase after being primed with it.

This is because the ADL phrase gives a (relatively) plausible interpretation of that stretch of audio. It is notable that a substantial number of listeners heard the ADL phrase even without the prime. In cold condition, 1% of listeners transcribed the full ADL phrase, and 25% transcribed the words *bank account*. Then, after additional listening, even in the PACT group, who were never exposed to the ADL prime, 13% gave the full ADL phrase in Section 2, while 19% of transcribers gave *bank account*.

By contrast, no participants heard the PACT phrase, in full or in part, in cold condition, or when primed with the ADL phrase. Indeed, of many hundreds who have heard this audio during previous research, none has ever heard anything remotely like the PACT phrase unless it has first been suggested to them – and even then, its priming effect is relatively small.

While these observations could be taken as confirming the validity of the ADL phrase (a suggestion we consider and reject below), their most important message is in confirming previous demonstrations² of the implausibility of *At the start we made a pact* – the phrase that played a crucial role in obtaining a murder conviction in the trial referred to above.

Surely, a system that allows manifestly implausible transcripts to be admitted as ‘assistance’ to a jury in evaluating indistinct covert recordings should be a cause for concern.

Other priming effects

Priming effects typically go beyond the interpretation of the phrase itself³. One way to quantify this, using data already presented, is to look at the number of *bank* transcripts given in Section 1 under different conditions (summarized in Figure 4). Recall that Section 1 is the part that supports (to the extent it does) the PACT phrase, not the ADL phrase, which includes the word *bank*.

Under cold condition, *back* is the most frequent interpretation, given by 13/76 (17%) of participants (mostly in a phrase such as *way back* or *made it back*), while *bank* is given by only 2/76 participants (3%). With the PACT prime, the proportion of *back* transcripts declines to 10%, while *pact* increases to 6% (of course a far larger percentage hear this as part of the full PACT phrase, not shown in Figure 4, but see Figure 3 above).

With the ADL prime, *back* retains the same percentage (17% – although note that this reflects different individual transcribers), while *bank* increases markedly to 37%. It might be worth emphasizing here that this *bank* is not a direct response to the ADL prime (which transcribers gave in Section 2) but to a general feeling, encouraged by their hearing of the word *money* and the ADL phrase that this audio as a whole has something to do with banks. Indeed a number of participants stated explicitly in comments that now they knew the audio was about an Adelaide bank account they could hear more than they had in cold condition.

Maybe he really did say Adelaide bank account?

For many non-experts in forensic phonetics, the fact that so many transcribe the full ADL phrase indicates it is a reliable transcript. However, this shows a misunderstanding of how human speech perception works. The fact that people ‘hear’ certain words, while it provides useful information, is far from definitive confirmation that those words were spoken.

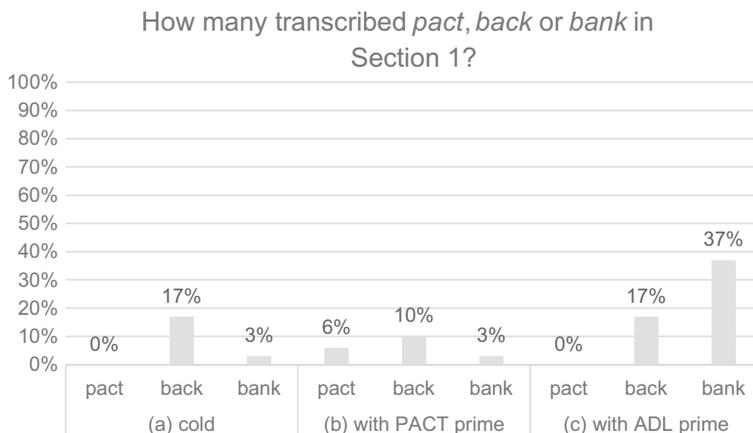


Figure 4. The number of participants who transcribed the words *pact*, *back* or *bank* under the three priming conditions. See text for details.

Consider the recent ‘swearing Minions’ scandal (readily searchable on the internet). Unfortunately for the producers of Minions (small yellow toys playing recorded nonsense prattle to amuse children), some parents heard the prattle as a string of obscenities. YouTube videos featuring the recordings, with a transcript, soon went viral. It is certain that testing the audio in an experiment like the one reported here would have yielded results at least as high as the ADL phrase.

In the Minions case, it was easy to accept the songs really were nonsense, and the rude words were in the ears of the beholders, created by a kind of pareidolia (the perceptual process that makes people see animals in the clouds, or faces in buttered toast).

But what if the Minions’ prattle had been evidence in a murder trial, and the swearwords had been relevant to the verdict? Would letting the decision as to the reliability of those YouTube transcripts be a matter of whether the jury ‘heard’ the suggested words give the right answer?

All speech, especially indistinct speech in poor quality recordings, is open to multiple interpretations. Deciding whether particular words were really spoken involves more than just crowdsourcing opinions about a transcript. Even phonetic analysis of the acoustics, although often helpful, does not always yield a definitive answer. Indeed, acoustic analysis performed without careful management of priming is prone to the kinds of unconscious cognitive bias that affects all other forms of forensic analysis⁴. This is one reason to be cautious of forensic opinions provided by audio engineers comparing non-experts’ transcripts to see which gains better support from their acoustic measurements

So – did he in fact say *Adelaide bank account*? What listeners are responding to in hearing this phrase is a sequence of syllables emerging from a longer stream of indistinct whispered speech in a noisy background. The overall rhythmic pattern of these syllables can be vaguely suggestive of a phrase such as *Adelaide bank account*. However, that interpretation takes no account of other syllables before, between or after the ones supporting that phrase.

This is evident from the fact that 23 of the 24 ADL transcribers who gave this phrase after priming gave it as an isolated phrase. Only one provided any words as context – and these words (*that Steve/she contributed to*) do not fit the acoustics, and are clearly influenced by the semantics. It is further supported by the observation that the full ADL phrase is more commonly heard than the part phrase *bank account*, while the word *bank* on its own is not heard at all in Section 2. Finally, informal testing suggests similar results could have been obtained for the word *barbecue* (‘heard’ by several participants, see Table 1).

The mark of a real expert is to know when to say ‘I don’t know’, and the truth is we simply do not know what the speaker said at this point in the recording.

Why didn’t the detective hear the ADL phrase?

As one last observation, it is worth reflecting on the fact that the detective who provided the transcript in the murder trial did not give the ADL phrase. Clearly the acoustics at this point are somewhat open to this interpretation and, interestingly, the case being tried did involve some connection with Adelaide, as well as with money and banks. It would have been quite easy, and quite useful to the prosecution, for the detective to have ‘heard’ these words. As it was, however, he was evidently (from his transcription of the PACT phrase a few seconds earlier) focused on ‘hearing’ something about a pact, in support of the case that the murder was a joint criminal enterprise. And indeed he transcribed part of the ADL section as a phrase making further reference to the pact he believed had just been mentioned.

Conclusion

Upon reading results like these, some are inclined to conclude ‘you can make people hear whatever you want them to hear’. That is not quite true. Under some conditions, listeners can be highly discerning. What is true is that speech perception is a complex interaction between a listener and an acoustic signal, whose outcome depends on the nature of the signal, the nature of the listener and the conditions under which the listening takes place. A good deal is known about these factors in phonetics and linguistics^{5–7}, with forensic research currently adding more, sometimes surprising, knowledge.

It is hoped that the small experiment reported here will add useful weight to two conclusions already well supplied with evidence^{8–11}, but sadly not yet provoking needed change in the legal process for handling indistinct covert recordings.

First, in general, evaluation of indistinct covert recordings is not a matter for ‘common knowledge’. If poor quality audio is to be admitted as evidence at all, it should only be with a transcript produced via a process that can ensure its reliability and suitability for purpose (along with careful advice and oversight regarding how the transcript should be used). For a range of reasons only touched on here, investigators’ transcripts, useful as they might be during investigations, will never meet that evidentiary criterion¹².

Second, more specifically, the police transcript that ‘assisted’ the court’s hearing of the covert recording in the murder trial it appeared in is thoroughly unreliable, yet nevertheless highly influential on listeners. The fact that transcripts by this particular detective have ‘assisted’ juries to reach guilty verdicts in at least two other murder trials might constitute a call to action for some readers (see further discussion in *Guest Editorial* this issue).

Acknowledgements

Many thanks to the participants who generously gave their time and effort to working through this experiment. Thanks also to Professor Kate Burridge for bringing the Minions to my attention, and to Professor Stephen Cordner for helpful comments on a prepublication version of this article.

Disclosure statement

No potential conflict of interest was reported by the author.

References

1. Fraser H, Stevenson B. ‘The power and persistence of contextual priming: more risks in using police transcripts to aid jurors’ perception of poor quality covert recordings. *Int J Evid Proof*. 2014;18:205–229.
2. Ibid.
3. Fraser H, Stevenson B, Marks T. Interpretation of a crisis call: persistence of a primed perception of a disputed utterance. *Int J Speech Lang Law*. 2011;18(2):261–292.
4. Thompson WC. What role should investigative facts play in the evaluation of scientific evidence?. *Aust J Forensic Sci*. 2011;43(2–3):123–134.
5. Ladefoged P, Disner SF. Vowels and consonants: an introduction to the sounds of language. 3rd ed. Oxford: Wiley-Blackwell; 2012.
6. Byrd D, Mintz TH. *Discovering speech, words and mind*. Oxford: Wiley-Blackwell; 2010.
7. Kreiman J, Sidtis D. *Foundations of voice studies*. Oxford: Wiley Blackwell; 2011.

8. Fishman CS. Recordings, transcripts, and translations as evidence. *Washington Law Rev.* 2006;81(3):473–523.
9. Miller AE. Jury suggestibility: the misinformation effect and why courts should care about inaccuracies in transcripts that accompany recorded evidence. *Law Psychology Rev.* 2016;40:363–382.
10. French P, Harrison P. Investigative and evidential applications of forensic speech science. In: Heaton-Armstrong A, ed. *Witness testimony: psychological, investigative and evidential perspectives.* 2006; p. 247–262.
11. Lange ND, Thomas RP, Dana J, Dawes RM. Contextual biases in the interpretation of auditory evidence. *Law Human Behavior.* 2011;35(3):178–187.
12. Fraser H. Transcription of indistinct forensic recordings: problems and solutions from the perspective of phonetic science. *Language Law/Linguagem e Direito.* 2014;1:5–21. <http://ler.letras.up.pt/site/default.aspx?qry=id05id1444id2692&sum=sim>