100
Transcription of poor quality covert recordings used as evidence in criminal trials

by

Helen Fraser
Last reviewed August 2017

AUTHOR INFORMATION

[100.10] Author

Helen Fraser

Dr Helen Fraser is an expert in phonetics, specialising in cognitive aspects of speaking and writing. She studied linguistics and phonetics at Macquarie University and the University of Edinburgh, and taught at several universities in the UK and Ireland before taking a lecturing position at the University of New England, NSW, which she held for 18 years. She is now an independent consultant, focusing on applications such as second language pronunciation, literacy acquisition and forensic transcription.

Contact:

https://helenfraser.com.au

Email:
helen@helenfraser.com.au

Chapter 100 © H Fraser

Published with permission of the author.

[The next text page is 100-101]

CONTENTS

[The next text page is 100-201]

INTRODUCTION AND DEFINITIONS

_____

[100.100]   Problems and solutions

Advances in technology are making it ever easier to collect *covert recordings* (conversations captured on tape, or its digital equivalent, without the knowledge of the speakers). Legally obtained covert recordings potentially provide powerful evidence in criminal trials, seemingly allowing the jury to listen directly as the recorded speakers make admissions or give information they would not have been willing to give openly.

Of course, the jury is not really listening directly to the conversation as they would have heard it in its original context, but to a representation of the conversation, played back in the context of the trial. To the extent the recording can be accepted as a full and faithful representation, and the context of the original conversation is known, this distinction may seem inconsequential. The problem is that covert recordings frequently provide a representation that is far from full and faithful. By their nature, the circumstances in which they are obtained limit control over the recording conditions. The result is that many covert recordings are of very poor quality, to the extent that all or part of the original conversation may be barely intelligible without the aid of a transcript. It is these cases that are the topic of the present chapter.

Over recent decades, the law has developed processes and practices for handling poor quality covert recordings, and presenting them as evidence, intended to mitigate these limitations. However, this development has involved little engagement with the science of forensic transcription and some of the practices have less than ideal results, as detailed below. There may be a case for general reform in this area, but in the meantime there is certainly a need for more critical evaluation of poor quality forensic audio used as evidence in individual cases – ideally by the side presenting them but especially by the opposing side.

This chapter aims to provide some background about the science of forensic transcription that may be useful in this regard. It starts by defining the scope of forensic transcription, then provides a brief overview of current legal practice in Australia, followed by a discussion of transcription from the perspective of the scientific disciplines of linguistics and phonetics. This allows development of a simple six-factor framework within which problems in the interpretation of poor quality covert recordings can be understood, and some directions in which solutions to those problems can be sought.

[100.120]   Preliminary definitions

Forensic audio

Forensic audio is any kind of audio recording used as evidence in the investigation or trial of a criminal matter. Within this broad definition, several cross-classifying distinctions can be drawn.

First is the distinction between audio in general and audio containing speech (*speech recordings*). As a rule, audio in which the main interest is in questions about the sound as such

(for example, identification of the recording conditions or the nature of any background noises) is best handled by experts in audio engineering as a form of *digital evidence* (see Chapter 101, *Digital Evidence* at [101.100]ff, in this publication). On the other hand, recordings in which the evidence of interest is speech (what is said and/or who said it) are best handled by experts in relevant branches of *phonetics* (the science of speech), for reasons to be discussed.

Speech recordings used as evidence in court (*forensic recordings*) can be *overt* (obtained with the knowledge of the speakers) or *covert* (obtained without the knowledge of the speakers). Overt recordings, such as police interviews, are usually of better quality and generally offer greater certainty regarding the identity of the speakers and the context in which they are speaking than do covert recordings.

Covert recordings can be obtained via telephone intercept or via a concealed microphone (sometimes called *environmental* or *ambient* recordings). In general, modern telephone intercepts are of better quality and come with firmer external evidence (in the form of location, date/time, caller/receiver numbers and so on) than do environmental recordings.

Another important distinction concerns the quality of the audio. Forensic speech recordings are rarely of high *fidelity* (faithfulness as a representation of the original conversation). Even overt recordings are usually only fair in quality and often feature significant background noise and other issues. Covert recordings are far more variable. Some can be described as clear, meaning that most of the speech can be understood readily, in one sitting, by anyone with competence in the language. Some are *poor quality or unclear* recordings, which can be defined informally (more detailed discussion to follow) as those where comprehension takes considerable effort, especially for those who do not know the context of the original conversation. Some poor quality recordings can be further classified as *indistinct*. These are so poor that casual listeners are able to understand little or nothing of the content. Even with repeated listening, it is still difficult to hear the conversation, and multiple transcribers working independently are liable to produce significantly different versions of the content (some examples are linked from forensictranscription.com.au). Ensuring reliable transcription of indistinct recordings generally requires consultation with an expert in forensic phonetics, for reasons to be discussed.

Forensic transcription

Forensic transcription is the process of representing the contents of a forensic speech recording with written symbols. (In practice, transcription often includes attributing utterances to particular speakers. However, while such speaker attribution is crucial to many cases, in the current chapter as in general usage, the focus is on the content of what is said, not on who said it.) It is also the branch of *forensic phonetics* which investigates this process scientifically and develops principles for evaluating the reliability of forensic transcripts, especially of poor quality recordings.

Finally, it is useful to distinguish *transcription of what is said by whom in a speech recording* from *interpretation of what was meant by the speakers in the original conversation*. The former refers to the actual words spoken at the time of the recording. In an indistinct recording, different listeners may have different opinions about this, but there is a fact of the matter, potentially retrievable through auditory and acoustic analysis by an expert in phonetics. By contrast, interpretation of what the speakers' intentions were in uttering the words is not a factual matter. In general, while expert analysis may shed useful light in some instances, experts are not better placed to judge speaker intentions than ordinary listeners, especially those in the jury, who have maximal knowledge of all relevant circumstances of the case (much of which is not available to expert witnesses).

[The next text page is 100-301]

CURRENT LEGAL PRACTICE & PERSPECTIVE

[100.300] Current legal practice regarding covert speech recordings

The handling, admission and presentation of forensic speech recordings are governed by a range of legal instruments, subject to interpretation by the judge in individual cases. Detailed coverage is well beyond scope of this chapter (see [13.5.280] and [2.5.270] in Freckelton and Selby, 2017; Edmond and San Roque, 2009). Here we present a very brief overview, highlighting a few key concepts relevant to the present discussion. For conciseness we focus on the most common scenario, in which speech evidence is presented by the prosecution. However, the principles developed in this chapter apply regardless of which side tenders the speech evidence (see also Fraser, 2014).

Most, though by no means all, covert speech recordings are obtained on behalf of police investigating a crime after granting of an official warrant. The recordings may be provided to police with a full or partial transcript prepared by a professional transcriptionist. Investigators listen to relevant sections and act on any intelligence they uncover to assist in collection of evidence relating to the crime. In many cases no further use is made of the audio evidence itself, which is superseded by other evidence more appropriate to the trial.

In some cases, however, parts of the recording may be needed as evidence in their own right. Where the recording is clear, this is straightforward. Following admission by the judge, the recording is simply played in court, with the jury listening to it as they do to all the other evidence, to evaluate its significance in relation to their verdict.

Generally, clear audio is accompanied by a full transcript prepared by a professional transcriptionist. This transcript is considered to be an aid to the jury's memory, or an *aide memoire*. This concept evidently relates to legal debate in the 1980s regarding whether the jury should be allowed to see daily transcripts of court proceedings. Those against argued the transcript could potentially influence the jury, by allowing overemphasis on particular items of evidence, while those in favour argued that the transcript was merely a convenient aide-memoire assisting the jury to recall information already well understood.

Where covert audio to be used as evidence is of such poor quality that the jury would be unlikely to hear it clearly if it were simply played in court, presentation is governed by a range of principles going back to the ruling in *Butera v Director of Public Prosecutions (Vic)* (1987) 164 CLR 180; 30 A Crim R 417.

One key principle is that the audio may be accompanied by a transcript to help the jury hear what is said (not just to help them remember what they have already heard). Another is that the transcript may be produced by an *ad hoc expert*. This is someone who, though possessing no real expertise in audio or transcription in general, is considered to have specialised knowledge relevant to perception of a particular indistinct recording, on the grounds they have familiarity with the speakers and the context and/or have listened to the audio many times. This role is

typically filled by police investigating the case. (It may be worth distinguishing this situation from one that holds in some other jurisdictions in which police-based audio units employ experts in phonetics to transcribe indistinct audio for all prosecution cases. The Australian situation is different in two main ways. First, our ad hoc experts have no level of expertise in audio or phonetics. Second, they are detectives working directly on the investigation of the case being tried.)

Of course, a transcript by an ad hoc expert is not simply accepted without scrutiny. The opposing side is expected to check it carefully and may present an alternative interpretation of any particular parts or even oppose the transcript as a whole. If this happens, the judge may examine the material personally and can rule on which, if any, version(s) of the transcript to admit, or even to exclude the audio evidence. Finally, the jury are cautioned to evaluate the transcript critically.

Another key concept strongly emphasised in *Butera* is that the evidence is the audio itself, not the transcript, and that its interpretation is ultimately a matter for the jury. For this reason, the judge is required to caution the jury that the transcript is provided as an *aide memoire*. This term harks back to the usage discussed earlier, but it is clear that in this context the transcript is intended as an *aid to perception*. The caution asks the jury to use it only as an aid, relying on their own judgement to decide what is said in the recording and what the speakers might have meant.

These concepts and practices have become familiar to lawyers over nearly 30 years of use and are taken for granted in the legal process. However, from the perspective of the science of forensic transcription, they raise a number of problems. These can be better explained after some brief background on the concept of transcription, first in legal, then in scientific, perspective.

[100.320] Transcription in legal perspective

The concept of a *transcript* is very familiar in legal contexts via the long-standing practice of using transcripts as official records of court proceedings. Traditionally these were produced by stenographers taking short-hand notes during the trial and writing them up each evening into a fair copy. The term *transcribe* first referred specifically to the process of "writing across" from the notes to the fair copy. Now it has broadened to cover the entire process of preparing a word-by-word or *verbatim* representation of spoken discourse as opposed to other kinds of record such as a summary or minutes.

This semantic broadening has been reinforced by recent extension of the term *transcribe* to mean typing up the contents of an audio recording, where ability to replay sections at will and correct the transcript on the fly reduces the need for separate preliminary notes. This modern usage is now very familiar in legal contexts, following the introduction of electronic records of police interviews. Reforms of the 1990s, intended to reduce the risk of police being accused of *verballing* suspects with a *fabricated confession*, mean that nearly all police interviews are now recorded as audio and/or video (Dixon, 2008). These recordings are routinely accompanied by transcripts typed up by professional agencies.

Overall, these familiar kinds of transcripts are highly reliable. Of course, errors do occur, but these are usually simple mishearings or misspellings, caused by lack of familiarity with the context (more common with increasing moves away from the traditional scenario where the transcriptionist is present at the original proceedings). Generally, reliability can be taken for granted, to the extent the transcripts can be treated almost like a copy of the audio, albeit in a different medium, with great advantages for convenience of access. For these reasons, it is common to refer to "the" transcript, with little need to question who made it or via what process.

This familiarity with transcripts in the legal process gives a sense that the nature of transcription is well understood. Transcription of poor quality forensic audio is seen as

essentially the same kind of process – as shown, for example, by the extension of the term *aide memoire* to cover aid to perception – with the main difference lying in the poor quality of the recording, which is compensated for by use of an *ad hoc expert* rather than a professional transcriptionist.

However, from a scientific perspective, the transcripts familiar in the legal process are but one kind within a much broader field. It is well recognised that the term *transcript* refers not just to a single type of written record but to a whole category of very different kinds of document. Different contexts require different kinds of transcript. Choosing the right kind, and preparing and using it appropriately, is an important aspect of research methodology. From a scientific point of view, transcripts of indistinct covert recordings are very different from familiar transcripts of court/interview recordings, and need to be handled very differently to ensure reliable interpretation of their content for reasons discussed in the next section at [100.500].

[The next text page is 100-401]

TRANSCRIPTION IN SCIENTIFIC PERSPECTIVE

————————————

[100.500] Introduction

Scientific study of language and speech is undertaken through various branches of linguistics and phonetics. These are little-known sciences that suffer from many misconceptions about their nature, even by specialists in other fields. It is often assumed they merely add technical detail to the *common knowledge* about language and speech that is widely shared in the community as a whole. However, advanced study of these disciplines brings realisation that much common knowledge about language and speech is founded on misconceptions (Harris, 1981; Bauer and Trudgill, 1998; Hudson, 1981; Jackendoff, 2012). Thus the boundary between *common* and *specialised knowledge* is hard to draw in these fields. This means special expertise is needed not just to provide additional information about speech evidence (valuable as that can be) but also to recognise the limitations of widely held but inaccurate beliefs.

The topic of transcription is a good example. It is commonly believed by those with little practical experience of transcription that it is a simple secretarial task of writing down what is *there to be heard*, a little like school-room dictation, and that producing a reliable transcript requires little more skill than decent spelling ability. This is far from true. Transcription is a complex process and the reliability of the resulting product is affected by a range of factors.

This can be seen by considering first some issues regarding transcription of good quality speech recordings and then discussing the additional issues raised in transcription of poor quality material.

[100.520] Word-level transcripts

The "verbatim" transcript familiar in legal contexts is a type of *orthographic* transcription (speech represented using ordinary spelling). Orthographic transcripts are used for a range of research purposes, including analysis of court proceedings and police interviews (Heffer, 2005; Heydon, 2005; Eades, 2010). However, linguists studying these topics generally prefer to create their own transcripts rather than use actual court or interview records. This is because careful analysis shows that so-called "verbatim" transcripts are far from accurate at a word-by-word level, with words frequently omitted, added or tidied up in the service of legibility (Eades, 1996). Linguistic research usually requires "messy" detail to be included and represented systematically via spelling and layout conventions.

The point here is not that linguists' transcripts are better than legal transcripts in an objective sense. Indeed, technical transcripts would be quite inefficient for the purposes for which legal transcripts are used, since the additional detail comes at the cost not just of increased time required for the preparation of the transcript but also of reduced legibility to a general audience.

Rather the point is that accuracy of a transcript is not an objective concept. A transcript is a representation of the audio (which itself is a representation of the original speech), and, as such, is necessarily partial. Even linguists' transcripts vary considerably depending on their context and purpose (Edwards, 2008; Roberts, 1997; 2006). Thus while *inaccuracy* in a transcript is relatively easy to define, defining *accuracy* is problematic. More helpful concepts are *reliability* and *fitness for purpose*. The upshot is that a transcript is never "the" transcript. It can only ever be "a" transcript. Understanding who made it and under what circumstances is crucial for evaluating its reliability.

These observations hold even for clear recordings where perception seems straightforward (assuming shared competence in the relevant language variety), leaving little room for disagreement regarding the content. The situation is more complex when we consider informal conversational speech – the kind usually captured in covert recordings.

In recent decades, technological advances have made it possible to obtain good quality recordings of unmonitored conversational speech, enabling linguists to undertake detailed study of spontaneous conversation. One of the biggest challenges in this type of research is preparation of the transcripts (Sidnell and Stivers, 2012; Schiffrin et al, 2008). Spontaneous speech is extremely "messy", with restarts, overlaps and other features that make it difficult to hear each word and identify each speaker – even in a good quality recording. (For an interesting, scholarly and readable account of the nature of spontaneous speech, see Erard, 2007.)

Most important, it is widely recognised that even the most detailed transcription involves selection and interpretation, whether according to explicit or implicit criteria, of the features to be represented (Green et al, 1997). For this reason, there can be a very large degree of variation in different transcripts of the same material prepared by analysts focusing on different kinds of information (Jefferson, 2004; Wald, 1995).

Recognition of this fact brings recognition of the inevitable risk of *cognitive bias* in preparation of transcripts. Cognitive bias, or suggestibility, is the unconscious tendency of the human mind to perceive what it expects to perceive, even if the physical evidence for that percept is limited (Kahneman, 2011; Ridley et al, 2013). It is quite distinct from personal or moral bias, which is a conscious tendency to prefer something that benefits one's own position. Cognitive bias is a broad phenomenon, relevant to recognition and memory of most kinds of perceptual input, including sound in general (see Chapter 69 *Hearing and the perception of sound*, at [69.50] ff, in this publication). However, the way it operates in the perception and representation specifically of speech has particular qualities, as discussed further below.

Since cognitive bias operates below the level of conscious awareness, it cannot be overcome by a mere effort of will (Thompson, 2011), but must be controlled for via systematic processes – familiar from the "double blind" experiments demanded by rigorous scientific research (Lapadat and Lindsay, 2001). For this reason, linguists conducting research that relies upon transcripts routinely use an independent transcriber for at least part of the material and are bound to report the degree to which that transcriber's work conforms with that of the researcher.

[100.540] Phonetic transcription

Transcription in the phonetic sciences focuses on the level of speech sounds and pronunciation, and opens up a whole realm of surprising discoveries about the nature of speech.

Even when a recorded conversation seems perfectly clear, close observation shows that individual sounds, words, phrases, even whole sentences, are blurred together, to the extent they are often unintelligible when excised from the overall context. This is very difficult to

explain convincingly without audio demonstrations. However, perhaps an impression can be given via the analogy that the clearest of speech is more like very messy handwriting than like printed text. Most people are familiar with the idea that individual letters or words in messy handwriting (even one's own) may be open to multiple interpretations, or uninterpretable, when taken out of their original context (Fraser, 2003). Surprising as it may seem, the same is exactly true of speech (see demonstrations linked from forensictranscription.com.au).

Thus, while phonetic transcription is often considered by non-specialists to be a daunting matter of mastering a set of complicated and confusing symbols for individual speech sounds (phonemes), it is actually the conceptual issues that are most challenging. Speech is an extremely complex signal (Kreiman and Sidtis, 2011; Laver, 1994; Ladefoged and Disner, 2012). It is literally impossible to include every aspect in a transcript (Kerswill and Wright, 2008), so the characteristics to be represented must be selected in light of specialist understanding of the sound system of the language and the purposes of the analysis (Heselwood, 2013; Cox, 2012).

Among other developments, these observations about the nature of speech have led the phonetic sciences to radically revise commonly accepted ideas about human speech perception (Cutler, 2012; Fraser, 2003). The blurred-together nature of speech makes it impossible that words and sentences could be recognised letter by letter, or *bottom up* (using only information available in the speech signal). Rather, perception involves a massive but unconscious *top down* component (use of the listener's learned knowledge about the language and the context in which the speech is interpreted). We say that context *primes* listeners' perception (sets up subconscious expectations that the incoming speech signal merely confirms or refutes). It may be useful to emphasise that priming is not in itself a bad thing. In fact it is essential. Speech perception, like other forms of perception, would be impossible without it. It does mean there is a great deal of very skilful but ultimately subjective top down processing going on behind the scenes of apparently effortless speech perception (Fraser, 2011; 2003).

The most interesting thing about this top-down processing is that listeners are wholly unaware of it, believing their perception to be a mere pick-up of information from "out there". Though this is a general feature of all cognition (Frith, 2007), it is both stronger and less recognised in speech than in other modalities (Jackendoff, 2012).

[100.560]   Poor quality recordings

Poor quality recordings are discarded for most scientific purposes. However, they have been the focus of longstanding research in their own right (Miller, 1951; Warren, 1999). One of the strongest findings is that, as bottom-up information in speech decreases (in other words, as the acoustic signal becomes less clear), the operation of top-down interpretation necessarily increases. Effectively, the listener's expectations, derived unconsciously from contextual knowledge or assumptions, play a greater role in perception than the (poor quality) information in the signal itself. Yet listeners continue to attribute their perception to the signal, unaware of the substantial contribution from their own cognitive processes.

Thus, even if perception is accurate, it can involve substantial 'over-interpretation' of the acoustics (attributing information to the recording which really comes from the contextual knowledge or assumptions of the listener).

The problem is that, to untrained listeners, the experience is the same whether valid contextual knowledge is helping them to hear accurately or invalid assumptions are creating perceptual error. In both cases, they feel confident they are simply "hearing what is there to be heard" and tend to accept their perception uncritically, even when alerted to the possibility of being primed (Bonifaz, 2014).

Since this phenomenon is rather little known outside phonetic science, it can seem hard to believe when first described. However, it is a very well established feature of human speech perception, with strong experimental support going back at least to the 1950s. A nice example is given by one of the experiments that first brought this interesting feature of human speech perception to light. Bruce (1958), investigating the effect of listeners' anticipations, created a number of sentences with the following form:

EXAMPLE 1

Sentence 1: I tell you that our team will win the cup next year.

Sentence 2: You said it would rain but the sun has come out now.

Participants heard these sentences "masked" with a hissing noise, which he calibrated so as to make the sentences around 25% intelligible. Next, he presented the same sentences in the same level of masking noise – but this time he preceded each with a keyword that gave it a context. For example, the keyword for Sentence 1 was SPORT, for Sentence 2, WEATHER and so on.

As predicted, sentences preceded by their keyword were more intelligible. However, an unexpected discovery was what happened when he played the masked sentences with the wrong keyword. This did not hinder perception, as had been predicted. Rather it created a different perception. For example, playing Sentence 1 with the keyword FOOD (instead of SPORT) led participants to hear a range of sentences quite different from the one that had actually been spoken, such as:

EXAMPLE 2

Sentence 1 (FOOD): I tell you that I feel more hungry than you are.

Playing the same Sentence 1 with the keyword TRAVEL led participants to hear yet another range of sentences, again different from the one that had actually been spoken, such as:

EXAMPLE 3

Sentence 1 (TRAVEL) I tell you that I too will leave next year.

Crucially, these erroneous perceptions were heard with no diminution of confidence. Listeners felt they were simply hearing what was there to be heard, in just the same way they did when their perception was accurate – an early indication of the now well accepted fact that listeners' personal confidence is a poor guide to the accuracy of their perception of indistinct audio.

[100.580] The science of forensic transcription

Transcription is a relatively small branch of forensic phonetics, where more attention has been given to speaker identification, or as it is more appropriately known, *speaker comparison* (Rose, 2007; Foulkes and French, 2012; Chapter 99, *Forensic Voice Comparison*, see [99.20]ff, in this publication) – though it is worth noting that the speaker attribution provided as part of the transcript can have significant influence on identification of voices (see [100.120]).

Even within forensic transcription, most research to date has focused on resolution of *disputed utterances* (French, 1990; Shuy, 1993; French and Stevens, 2013; Coulthard and Johnson,

2007). This is the situation where there is a specific dispute regarding what is said in a particular section of a forensic speech recording. Auditory and acoustic analysis by a qualified phonetician can be of considerable assistance in resolving such disputes.

Traditionally, conclusions of such analyses have been expressed as support for one or other of the suggested interpretations, or proposal of another, with appropriate levels of confidence. Unfortunately, if several experts give different low-confidence proposals, this can sometimes be interpreted by the courts as an invitation to let the judge, or the jury, choose which one they prefer, which was not necessarily the experts' intention. For this reason, there has recently been some recognition of the value, where phonetic analysis does not allow transcription with a level of certainty appropriate for forensic evidence, of experts declaring an utterance *untranscribable*. This gives the evidence a similar status to a smudged fingerprint, or inconclusive DNA analysis, which is typically excluded from the trial, rather than being left to the court to evaluate.

Useful and important as disputed utterance analysis is, it does not cover all, or even most, of the problematic situations in forensic transcription. For one thing, it depends on the side opposing the transcript putting forward an alternative interpretation of particular phrases. Arguably, this should not be necessary, as the onus should be on the side producing the transcript to demonstrate its accuracy. In any case, it is not always possible, especially in the most typical scenario, where an entire indistinct speech recording is accompanied by a lengthy transcript, with no prior indication of which particular phrases might be relied upon in the trial.

It is clear that the scientific discoveries discussed above are relevant to this more typical scenario (Bucholtz, 2009), but it has received relatively little academic attention. This may be because it is not often brought to the attention of experts, being most commonly dealt with via the legal process outlined in [100.300]–[100.320], which considers speech perception to be a matter of common knowledge (Shuy, 2009). This lack of engagement with the science of transcription has led to a range of problems with the handling of indistinct covert recordings in the legal process, to which we turn next.

PROBLEMS WITH CURRENT LEGAL PRACTICE REGARDING INDISTINCT COVERT RECORDINGS

_____

[100.700]   Inaccurate transcripts

Current legal practice (see [100.500]–[100.560]) depends heavily on transcription by detectives given the role of so-called *ad hoc expert* on the basis of specialised knowledge they supposedly derive from listening many times to the audio. This is problematic in several respects. First, repeated listening is no guarantee of accuracy (if it were, it would surely be more cost effective to pay professional transcriptionists to listen many times than to have detectives do this work). In fact, it is quite possible to listen many times and be wrong every time (Fraser and Stevenson, 2014).

Second, whatever advantage detectives may have in interpreting indistinct audio relating to their cases, it comes not from repeated listening but from their knowledge of the context in which the recording was obtained. It is important to acknowledge that listeners with contextual knowledge can potentially interpret parts of a recording that are unintelligible to those without such knowledge. For this reason, detectives' contextual knowledge can and should be harnessed in interpreting indistinct recordings for investigative purposes (as discussed further below; see also Fraser (2014)). The problem in using their transcripts in court, without scrutiny by a genuine, and independent, expert, is that contextual knowledge is a double-edged sword. As described above, it can cause listeners to hear confidently but inaccurately: see [100.580]. For this reason, its use must be managed very carefully, in ways of which ad hoc experts have no understanding.

Finally, even to the extent ad hoc experts' contextual knowledge may help them understand particular phrases within an indistinct recording, it does not help them produce a reliable transcript of the recording as a whole. For these and other reasons, police transcripts are frequently inaccurate as well as suffering from a range of other problems (see further discussion in Fraser, 2014).

[100.720]   Inadequate checking

In the current system, once an ad hoc expert's transcript is in circulation, it is rare for anyone to listen to the audio without being *primed* by that transcript (see [100.540]). The opposing side does check the transcript, but this is typically done by lawyers, or the defendant, simply reading the transcript as they listen to the audio in real time. While in some cases, this may result in identification of one or more disputed utterances, it is not an effective way to evaluate transcripts of poor quality audio, since seeing a transcript can prime the perception even of those who might seem predisposed to disagree with it. Indeed, it may produce a "curse of knowledge" effect, whereby listeners feel the audio is far clearer than it really is (Lange et al, 2011).

For this reason, it is alarmingly common for manifestly inaccurate transcripts to be admitted as a resource to "assist" the jury's perception of indistinct covert recordings (see forensictranscription.com.au). While this is sometimes discovered during the trial, there are numerous cases where it has not been picked up till after the trial is concluded, and potentially many others where it might not have been noticed at all.

Audio from one such case was used as the basis of an experiment (Fraser and Stevenson, 2014) aiming to shed light on these issues. The experiment was conducted in two parts. The first part demonstrated that, in the absence of contextual knowledge, the police transcript was not just inaccurate but actually quite implausible. Its priming effect was less than usual when first presented to participants, and even the few who initially accepted it, quickly abandoned it when presented with a more plausible alternative.

This raised the question of how the transcript could ever have been accepted in court. An answer was suggested by the second part of the experiment. This began by giving a new group of participants contextual knowledge about the trial, similar to the background that had been available to lawyers on both sides of the case, and, later, to the jury. These participants were far more likely than those in the first part to hear the audio in line with the inaccurate police transcript and far less likely to be swayed by the alternative, more plausible, transcript. As well, many subtle effects on perception were found, even of participants who disagreed with the police transcript.

Importantly, before-and-after questions revealed a strong effect of seeing the inaccurate transcript on participants' opinions about the guilt of the defendant, even in those who rejected its exact words. In short, participants believed their understanding of the case was influenced by the audio, not realising how much their perception of the audio had been unconsciously influenced by their understanding of the case.

These and other results were interpreted as evidence that leaving evaluation of forensic transcripts to defendants and their legal representatives gives insufficient protection from inaccuracy.

[100.740] Unrealistic expectations of jury

The jury have the important role of interpreting the significance of the original conversation captured in the forensic speech recording in relation to all the other evidence presented at the trial. This involves evaluating the intentions of the speakers, their truthfulness, the genuineness of their emotions, the presuppositions and implications of what they say, and so on. To make this interpretation, the jury need to pay careful, unbiased, attention to features such as the tone of voice of the speakers in the recording, which, in turn, requires clarity and confidence regarding what words are being said (see [100.300]–[100.320]).

The current system intends to provide this clarity by giving them a transcript prepared by an ad hoc expert and cautioning them to use it only as an aid. Unfortunately, what the caution asks of juries is not possible for them to do, for reasons highlighted by another recent experiment (Fraser et al, 2011). This used the disputed utterance from a real case and closely simulated the experience of a jury attempting to evaluate it against a demonstrably inaccurate police transcript.

Subjects, divided into two experimental groups, listened repeatedly to the same audio, while evidence about the case, including the inaccurate police transcript, was gradually revealed to them across seven evidence points. At each evidence point, they were asked what they heard in the audio, along with various other questions.

At first, virtually no one in either group heard anything remotely like the police transcript. However, at Evidence Point 4, when the inaccurate transcript was explicitly suggested, over

30% of Group A, who were given the inaccurate police transcript, confidently heard the exact words of the transcript, with many others displaying perception clearly influenced by the transcript.

Subsequent evidence points simulated attempts to help participants recognise that the police transcript was indeed inaccurate. However, these brought about little change in perception. The final evidence point explained the purpose of the experiment, with information that the materials had been deliberately chosen to show how easily a demonstrably inaccurate transcript could mislead listeners' perception of indistinct audio. However, even after being told that experts on both sides of the case had agreed the police transcript was inaccurate, 17% of Group A claimed to hear its exact words, with, again, many more influenced by it in a variety of ways, and, again, a strong effect of seeing the inaccurate transcript on participants' ultimate opinions about the guilt of the defendant, even in those who rejected its exact words.

Most interesting of all was the response of Group B at this last evidence point. Group B had been primed with a different transcript and had never heard anything like the police version. However, mention of the (inaccurate) police transcript at the final evidence point prompted 12% of Group B to hear the audio exactly in line with the transcript, and many others to be influenced by it.

Finally, it is worth noting that participants in the experiments spent massively more time listening to individual phrases, with massively more encouragement to critique the transcript, than is possible for typical juries (Fraser and Stevenson, 2014).

These results and other considerations suggest that it is unrealistic to expect a jury to use a transcript "only as an aid" and to reach their own independent conclusion as to what is said in an indistinct recording. An inaccurate transcript will inevitably influence their perception of what words were spoken, and by whom. This in turn must affect their ability to form a valid interpretation of the speakers' intentions, and ultimately limit their ability to reach a fair verdict.

FACTORS AFFECTING THE RELIABILITY OF TRANSCRIPTS

[100.900] Creating a suitable framework

Finding solutions to complex problems like these requires a principled understanding of the factors that affect the reliability of a transcript. Since research relevant to forensic transcription is spread across several little-known branches of linguistics and phonetics, as discussed the previous section, it may be useful to distil the information into a framework suitable for discussion of issues in forensic transcription, and enable more detailed consideration of what makes speech recordings seem clear or indistinct (see [100.120]).

[100.920] Recording factors

Recording factors have to do with the technical nature of the recording, including aspects such as the quality and settings of the equipment used, the proximity of the speakers to the microphone, the "signal-to-noise ratio" and so on. All else being equal, a good recording is better than a poor one. However, recording factors interact strongly with other factors, to the extent it is rare that poor recording conditions are the only or main factor making audio hard to hear. Indeed, speech recorded with high production values can be quite indistinct (think of the dialogue in some modern movies). This is because recording factors have more influence on *listenability* (how easy or pleasant the recording is to listen to) than *intelligibility* (whether the words can be heard clearly). While a recording requires a certain minimal quality to be intelligible, above this baseline, most quality improvements affect only listenability. In covert recordings, of course, it is intelligibility that is the issue of concern.

[100.940] Speech factors

Speech factors are the characteristics of the speech in the original conversation. In formal situations, such as court proceedings or police interviews, speakers take separate turns so that only one voice is heard at a time and speech is monitored for clarity, whether by the speakers themselves or by an interviewer requiring them to "speak up for the tape". This kind of speech can often be understood even in a poor recording. Spontaneous conversation, on the other hand, even of a kind that is readily intelligible in its original context, can be hard to transcribe even in a good recording (due to overlaps, restarts and other features discussed earlier (see [100.520]) and very difficult in a poor one.

[100.960] Context factors

The importance of context in human speech perception has been emphasised. Much of what is called "context" is really the listener's knowledge or assumptions about the circumstances of the recording, discussed in the next section. Here we focus on the physical context of other speech surrounding the speech that is to be transcribed. It has long been known that meaningful sentences are more easily understood, even in a poor recording, than random words (Byrd and Mintz, 2010). This means that, in general, it is an advantage to have extended recordings with good continuity and cohesion (content that "hangs together", with internal cross references and a flow of topics). On the other hand, it can be very difficult to give a reliable transcript of speech that is heard only in short snatches, whether due to poor recording conditions or because the original speech itself consisted of isolated remarks.

[100.980] Listener factors

The first three factors, taken together, affect the overall *clarity* of the recording. Even more important, however, is the knowledge, or expectations, of the listener. Any transcriber is first a listener, so all the characteristics of human speech perception play a part (as discussed in [100.500]–[100.580]). For any recording, but especially for an indistinct one, it is essential to consider the role of *top down* processing and the power of *priming* (see [100.540]).

For this reason, it is useful to classify listeners as *independent* or *involved*. *Independent* listeners are those who, while they may have relevant knowledge of the context of the recording and purpose of the transcript, have no particular stake in the ultimate use of the transcript. Professional transcriptionists are carefully chosen so as to have this kind of independence. By contrast, *involved* listeners have some interest (whether conscious or not) in how the audio they are transcribing might ultimately be interpreted. Finally, it is important to recall the almost irresistible *priming* effect, of having seen another listener's transcript before hearing the audio (see [100.720]–[100.740]).

[100.1000] Transcriber factors

Having perceived the speech, the next step is to transcribe it. This is a far more demanding task than is recognised by those who have never tried it. Even formal, monitored speech in a good recording requires repeated, patient listening just to write it out word by word – and testing the result against an authoritative version shows how elusive true accuracy can be. As discussed above (see [100.540]), transcription of conversational speech is far harder. In addition, training is necessary to lay out the transcript in a way that makes it fit for the purpose required by the end user (see Fraser, 2014).

For these reasons it is useful to distinguish *professional* transcribers from those who are *untrained*. Professionals have typically been trained in how to lay out and present a transcript, and have usually undergone some kind of accuracy testing. It is worth emphasising that this kind of professional experience, valuable as it is, is not the same as genuine expertise in transcription, attained through high-level qualifications in relevant branches of linguistics and phonetics, such as those discussed in [100.500]–[100.580].

Untrained transcribers have neither professional experience nor real expertise. They may work very hard on a transcript, listening many times to the audio and updating it frequently to reflect changes in their hearing. However, due to their lack of experience with the variability of speech perception, they may be inclined simply to accept their most recent hearing as accurate. Further, they lack understanding of how to lay out a transcript to make it suitable for its intended purpose (see Fraser, 2014).

[100.1020] Situational factors

Situational factors relate to the overall situation in which the transcript is prepared and the purpose for which it is to be used. While reliability is affected by all the factors just discussed,

it can really only be assessed in relation to the situational factors. Most important of these is availability of *ground truth* (authoritative knowledge), or at least reliable *external evidence* (evidence from outside the speech recording itself), regarding what was said by whom in the original conversation, against which the transcript can be evaluated. Of course, in the case of indistinct covert recordings used as evidence in criminal trials, ground truth is lacking and external evidence is often limited in quantity and reliability. The question of how to evaluate transcripts in this situation is a crucial one, to which we return shortly.

[The next text page is 100-701]

USING THE FACTORS TO UNDERSTAND THE PROBLEMS IN FORENSIC TRANSCRIPTION

[100.1100]   Court transcripts v transcripts of covert recordings

As discussed in [100.320], the legal process has a tendency to assume transcription of indistinct covert recordings is a similar process to transcription of court or interview proceedings, except that the covert recordings have poor recording conditions. In other words, there is too much focus on recording factors. Considering all the factors set out above helps us to see this is far from true. The two situations are different on every factor.

Court/interview transcripts are based on fair quality recordings of formal, monitored speech with good cohesion. Listeners know, or can make valid assumptions about, the circumstances of the recording, but have no direct involvement in the outcome. Transcribers are independent professionals trained and tested in use of appropriate standards and conventions. There is a high level of certainty regarding what was actually said and by whom, with requirement that the transcript be checked by speakers and/or listeners shortly after the recording is made. In short, their reliability stems from accountability designed in to the system at every stage.

Indistinct covert recordings, by contrast, are poor quality recordings, typically of conversational speech and often with low cohesion. The context in which they are heard by the court is very different from the context in which they were originally spoken, which may be only partially known. Ad hoc experts are involved listeners and untrained transcribers. Ground truth, by definition, is not available, and external evidence is variable in availability and reliability.

This paradoxical situation, in which the worst audio is transcribed and checked by the least qualified personnel, and used for purposes with the highest stakes, would appear to arise from poor appreciation of the complexity of transcription. Getting to grips with some of that complexity may hold the key to finding good solutions, and, perhaps even more importantly, avoiding bad ones.

[The next text page is 100-801]

TOWARDS SOME SOLUTIONS

What not to do

[100.1200] "Enhancing" does not improve intelligibility

One effect of focusing on *recording factors* (see [100.920]), to the neglect of other factors, is that it has become routine for "enhanced" versions of indistinct audio to be used for transcription and admitted as evidence, often with little scrutiny. Full discussion is beyond the scope of the current chapter (see forensictranscription.com.au). However, it is important to be clear that, despite the impression given in the media, there are currently no automatic processes that can make unintelligible audio intelligible (Loizou and Kim, 2011; Jessen, 2008).

In certain limited circumstances, it is possible to use specialised hand techniques to clarify short indistinct utterances – though great care must be taken to avoid unintentionally creating an auditory impression different from the original speech.

However, the "enhancing" routinely carried out on forensic audio uses techniques that were originally designed to improve *listenability* (see [100.240]). They have at best negligible, and sometimes negative, effects on intelligibility. However, being informed that audio has been "enhanced" can increase listeners' confidence in perception primed by an inaccurate transcript. For these and other reasons, the term should be avoided, and the effectiveness of "enhancing" techniques should be scrutinised with far more rigour than is currently common.

[100.1220] Presenting audio with no transcript does not avoid priming

The problems of priming are not solved by letting the jury hear the audio without any transcript, as *contextual priming* is at least as powerful as *priming by a transcript* (Fraser and Stevenson, 2014).

[100.1260] Transcripts made with no knowledge of context are not more accurate

Problems of priming are not solved by completely withholding context from the transcriber. For reasons similar to those discussed in relation to disputed utterances (see [100.580]), the practice of obtaining transcripts from multiple transcribers for the court to choose among is

problematic. Especially with indistinct audio, appropriate knowledge of context is essential to reliable transcription. The important thing is to manage the disclosure of context so that *listener factors* can be kept separate from other factors, and to ensure the transcriber is *independent* rather than *involved* in the case.

[100.1280] Personal evaluation of audio by the judge is not the solution

There have now been several cases in which judges have listened personally to indistinct audio to make a ruling on the effectiveness of "enhancing", or the validity of a transcript. While this is certainly a step in the right direction, and has been useful in raising awareness of issues surrounding covert recordings, it is not appropriate as a long-term solution. This is for a range of reasons, including the time frequent resort to this strategy would demand of the judge. Not least of these reasons, however, is that judges are not significantly less prone to priming than anyone else (as demonstrated by two recent experiments, see forensictranscription.com.au). What is really needed is recognition that phonetics is a science like other sciences, and development, in close collaboration between phonetic science and the law, of a uniform, accountable process for the handling of all covert recordings. More will be said about this below.

[100.1300] Technical linguistic/phonetic transcripts are not suitable for juries

Perhaps surprisingly, it is not being recommended that all forensic recordings should be accompanied by technical transcripts produced by genuine experts in transcription. This is partly because the cost would be prohibitive. A more important reason, however, is that most of the technical types of transcript used in linguistics and phonetics are not suitable for the purpose transcripts are used for in court – that is, to give the jury reliable certainty about the words so they can focus on tone of voice and other features that, combined with other evidence presented in the trial, let them reach their own evaluation of the speakers' intentions. This is a very different situation from the usual purpose of transcripts in linguistics and phonetics, and means that forensic transcription (like other branches of forensic science) is not a mere *application* of existing theoretical knowledge, but requires development of theoretical knowledge in its own right.

There is an important role for genuine experts in forensic transcription. However, it is not primarily the role of producing transcripts (though expert transcripts may be appropriate in some situations). Rather it is the role of designing, in collaboration with lawyers, an accountable process for the handling of indistinct covert speech recordings used as evidence in criminal trials.

An accountable process for handling indistinct covert speech recordings

[100.1320] Adopting a different approach

Court/interview transcripts are reliable because they are produced via a process with many steps, each with appropriate checks and balances. Transcription of poor quality covert recordings needs a process of similar accountability. However, due to the very different combination of factors affecting their reliability, the nature of the process, and of the checks and balances, will be different.

One of the major goals of such a system would be to exclude *untranscribable* recordings at an early stage to avoid the wasted effort of arguing about their content. Another would be to

distinguish rigorously between interpretations of covert recordings used purely for *investigative* purposes and those used in court for *evidentiary* purposes (see further discussion in Fraser, 2014). Most important would be to develop an efficient, methodical and standardised process for ensuring that all covert recordings used as evidence in court are accompanied by a reliable transcript. Hopefully admission of covert recordings could then include scrutiny of the transcript as evidence in its own right, recognising that it will do far more than "assist" the jury's perception of the audio evidence. (See arguments in favour of such scrutiny from a legal perspective, from the United States, in Fishman, 2006.)

Designing and implementing such a process is a major, though surely worthwhile, undertaking requiring collaborative research between experts in transcription and in law. In the meantime, the discussion of this chapter suggests some simple recommendations that can be used on a case by case basis.

Recommendations for case work

[100.1340]   Practical steps

From the perspective of phonetic science, if poor quality audio is to be used as evidence, it should be accompanied by a *reliable* transcript that removes from the jury the responsibility of deciding what words are spoken in the recording and allows them to concentrate on interpreting what the speakers meant by the words (see [100.120]).

Transcripts by "ad hoc experts" are not suitable for this purpose, for all the reasons discussed above. However, they will no doubt continue to be used for the foreseeable future. This suggests a need for increased diligence in assessing their reliability, ideally by the side presenting them, but certainly by the opposing side.

The best advice is to avoid "checking" by having someone involved in the case listen to the audio while reading the ad hoc expert's transcript. This brings the strong possibility of *priming*, to the extent that significant errors or over-interpretations in the ad hoc expert's transcript are liable to go unnoticed – even when those doing the checking feel they are paying very close attention (see [100.700]–[100.1020]).

A better way of checking an ad hoc expert's transcript is to compare it with a transcript prepared by an independent professional transcriptionist who has no knowledge of the context (see [100.980]–[100.1020]). In many cases, such a transcript will already exist, as police transcripts often add to previous versions produced by an agency. If no existing independent version is available, one should be prepared.

This independent transcript will not give a definitive interpretation of the audio. However, comparing it with the ad hoc expert's transcript might provide a useful impression of whether the audio should be treated as *clear*, *unclear* or *indistinct* (see [100.120]). If the independent transcriber hears the audio substantially differently from the ad hoc expert, this indicates the audio is *indistinct*, and suggests the ad hoc expert's transcript has been strongly influenced by subjective *top-down* processing. This does not necessarily mean the ad hoc expert's transcript is wholly inaccurate. However, from the perspective of phonetic science, it suggests a need for the side wishing to use the transcript to obtain the opinion of a genuine expert to back up the ad hoc expert's version. From the legal perspective, there may still be cases where it is considered valid to let evaluation of the ad hoc expert's transcript be a matter for the jury. In such cases, the burden may fall on the opposing side to obtain an expert evaluation of those parts of the ad hoc expert's transcript that are most relevant to the trial.

[The next text page is 100-901]

CONCLUSION

[100.1500] Short-term and long-term changes

It is not too strong to suggest that the current system for handling poor quality speech recordings used as evidence in criminal trials brings back the potential for *verballing* (see [100.320]) that was supposed to be eliminated by the introduction of processes for recording and transcribing police interviews. A full solution to this problem requires a similar level of change at the system level to the one involved in introducing the requirement for electronically recorded interviews – and it may be worth noting that this change, though disruptive at the time, is now universally welcomed by all participants in the legal process.

In the meantime, heightened awareness on the part of individual prosecution and defence teams, and of judicial officers, can bring about improvement. It is hoped that this chapter might help in that regard.

BIBLIOGRAPHY

[100.1600] References

See also forensictranscription.com.au for more detailed discussion of all points, including audio demonstrations, description of experimental results and case studies.

Bauer L and Trudgill P, *Language Myths* (1998) ePenguin

Bonifaz S, *The Effect of Priming Awareness when Listening to Disputed Utterances* (MSc Dissertation, Department of Language and Linguistic Science, University of York, 2014).

Bruce DJ, "The Effect of Listeners' Anticipations on the Intelligibility of Heard Speech" (1958) 1 *Language and Speech* 79–97.

Bucholtz M, "Captured on Tape: Professional Hearing and Competing Entextualizations in the Criminal Justice System" (2009) 29 (5) *Text and Talk - An Interdisciplinary Journal of Language, Discourse & Communication Studies* 503–523.

Byrd D and Mintz TH, *Discovering Speech, Words and Mind* (Wiley-Blackwell Publishing, 2010).

Coulthard M and Johnson A, *An Introduction to Forensic Linguistics: Language in Evidence* (Routledge, London and New York, 2007).

Cox F, *Australian English Pronunciation and Transcription* (Cambridge University Press, 2012).

Cutler A, *Native Listening* (MIT Press, 2012).

Dixon D, *Videotaping Police Interrogation* (University of New South Wales Faculty of Law Research Series 28, 2008).

Eades D, "Verbatim Courtroom Transcripts and Discourse Analysis" in Kniffke H (ed), *Recent Developments in Forensic Linguistics* (Peter Lang, Frankfurt, 1996), pp 241-254.

Eades D, *Sociolinguistics and the Legal Process* (Multilingual Matters, 2010).

Edmond G and San Roque M, "Quasi-Justice: Ad Hoc Expertise and Identification Evidence" (2009) *Criminal Law Journal* 338–33.

Edwards JA, "The Transcription of Discourse" in Deborah Schiffrin et al (eds), *The Handbook of Discourse Analysis* (Blackwell Publishing Ltd, Oxford, UK, 2008) pp 321–348.

Erard M, *Um: Slips, Stumbles and Verbal Blunders, and What They Mean* (Random House, New York, 2007).

Fishman CS, "Recordings, Transcripts, and Translations as Evidence" (2006) 81 *Washington Law Review* 473–523.

Foulkes P and French P, "Forensic Speaker Comparison: A Linguistic-acoustic Perspective", in *Oxford Handbook of Language and Law* (Oxford, Oxford University Press, 2012), pp 557–572

Fraser H, "Issues in Transcription: Factors Affecting the Reliability of Transcripts as Evidence in Legal Cases" (2003) 10(2) *International Journal of Speech Language and the Law* 203-226.

Fraser H, "Spoken Word Recognition" in Chapelle CA (ed), *Encyclopedia of Applied Linguistics* (Wiley-Blackwell, 2011).

Fraser H, "Transcription of Indistinct Forensic Recordings: Problems and Solutions from the Perspective of Phonetic Science" (2015) 1(2) *Language and Law/Linguagem e Direito* 5-21.

Fraser H and Stevenson B, "The Power and Persistence of Contextual Priming: More Risks in Using Police Transcripts to Aid Jurors' Perception of Poor Quality Covert Recordings" (2014) 18 *International Journal of Evidence and Proof* 205-229.

Fraser H et al, "Interpretation of a Crisis Call: Persistence of a Primed Perception of a Disputed Utterance" (2011) 18(2) *International Journal of Speech Language and the Law* 261-292

Freckelton I and Selby H, "Police Evidence" in *Expert Evidence* (Thomson Reuters, Australia, 2017).

French P, "Analytic Procedures for the Determination of Disputed Utterances" in Kniffke H (ed), *Texte zu Theorie und Praxis forensischer Linguistik* (Tübingen, Niemeyer, 1990) pp 201–213.

French P and Stevens L, "Forensic Speech Science" in Knight R and Jones MJ (eds), *Bloomsbury Companion to Phonetics* (Continuum International Publishing Group, London, 2013).

Frith C, *Making Up the Mind: How the Brain Creates Our Mental World* (Wiley-Blackwell, 2007).

Green J et al, "The Myth of the Objective Transcript: Transcribing as a Situated Act" (1997) 31(1) *TESOL Quarterly* 172–176.

Harris R, *The Language Myth* (Duckworth, London, 1981).

Heffer CN, *The Language of Jury Trial: A Corpus-Aided Analysis of Legal Discourse* (Basingstoke, Palgrave MacMillan, 2005).

Heselwood B, *Phonetic Transcription in Theory and Practice* (Edinburgh University Press, Edinburgh, 2013).

Heydon G, *The Language of Police Interviewing: A Critical Analysis* (Palgrave MacMillan, Houndmills, 2005).

Hudson R, "Some issues on which linguists can agree" (1981) 17 *Journal of Linguistics* 333-343.

Jackendoff R, *A User's Guide to Thought and Meaning* (Oxford University Press, Oxford, 2012).

Jefferson G, "Glossary of Transcript Symbols with an Introduction", in Lerner GH (ed), *Conversation Analysis: Studies From the First Generation* (John Benjamins, Amsterdam, 2004).

Jessen M, "Forensic Phonetics" (2008) 2(4) *Language and Linguistics Compass* 671-711.

Kahneman D, *Thinking, Fast and Slow* (Farrar Straus Giroux, New York, 2011).

Kerswill P and Wright S, "The Validity of Phonetic Transcription: Limitations of a Sociolinguistic Research Tool" (2008) 2(3) *Language Variation and Change* 255-275.

Kreiman J and Sidtis D, *Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception* (Wiley Blackwell, Oxford, 2011).

Ladefoged P and Disner SF, *Vowels and Consonants: An Introduction to the Sounds of Language*, 3rd ed (Wiley-Blackwell, 2012).

Lange ND et al, "Contextual Biases in the Interpretation of Auditory Evidence" (2011) 35(3) *Law and Human Behavior* 178-187.

Lapadat JC and Lindsay AC, *Examining Transcription: A Theory-Laden Methodology*, Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA, April 13-17, 1998, 1-20.

Laver J, *Principles of Phonetics* (Cambridge University Press, Cambridge, 1994).

Loizou PC and Kim G, "Reasons Why Current Speech-Enhancement Algorithms Do Not Improve Speech Intelligibility and Suggested Solutions" (2011) 19(1) *IEEE Transactions on Audio, Speech, and Language Processing* 47-56.

Miller GA, *Language and Communication* (McGraw-Hill, New York, 1951).

Ridley AM et al, *Suggestibility in Legal Contexts* (Wiley-Blackwell, 2013).

Roberts C, "Issues in Transcribing Spoken Discourse" (2006) *Qualitative Research Methods*.

Roberts C, "Transcribing Talk: Issues of Representation" (1997) 31(1) *TESOL Quarterly* 167-172.

Rose P, *Forensic Speaker Identification* (Taylor & Francis, 2007).

Schiffrin D et al (eds), *The Handbook of Discourse Analysis* (Blackwell Publishing Ltd, Oxford, 2008).

Shuy RW, "A Dozen Reasons Why Linguistic Expertise is Rejected in Court" (2007) http://www.rogershuy.com/slr_selected_articles.html, accessed 18 July 2017.

Shuy RW, *Language Crimes: The Use and Abuse of Language Evidence in the Courtroom* (Blackwell, Oxford, 1993).

Sidnell J and Stivers T, *The Handbook of Conversation Analysis* (John Wiley & Sons, 2012).

Thompson WC, "What Role Should Investigative Facts Play in the Evaluation of Scientific Evidence?" 43 (2-3) *Australian Journal of Forensic Sciences* 123–134.

Wald B, "The Problem of Scholarly Predisposition" (1995) 24(2) *Language in Society* 245–257.

Warren R, *Auditory Perception: A New Analysis and Synthesis*, 2nd ed (Cambridge University Press, Cambridge, 1999).

[The next text page is 101-1]