Automatic forensic voice comparison
(Automatic forensischer Stimmenvergleich)

Timo Becker

Forensic Researcher and Expert
Bundeskriminalamt
Forensic Science Institute
Department of Speaker Identification and Audio Analysis (KT54)
Wiesbaden, Germany

Awarding Institution: Universität Trier, Germany
Date of Award: 2012

KEYWORDS FORENSIC VOICE COMPARISON, AUTOMATIC SPEAKER RECOGNITION,
FORMANT ANALYSIS, AUTHENTIC FORENSIC AUDIO RECORDINGS

Automatic speaker recognition systems have been developed and constantly improved during recent decades. However, even though these systems have been applied to the task of forensic voice comparison since the early 2000s, their behaviour under realistic forensic conditions has not often been thoroughly tested since then. Additionally, the application of these systems is not standardised and is far from being widely accepted in the forensic community. In order to provide a realistic estimate of the performance of automatic forensic voice comparison systems, two approaches to automatic forensic voice comparison developed by the author are tested under forensic conditions in this...
thesis. Additionally, these systems are compared with two established automatic forensic voice comparison systems that are already in use in forensic laboratories. The author investigates the performance of automatic forensic voice comparison in order to answer the question whether such systems are able to provide useful information.

Automatic forensic voice comparison systems generally compute a similarity measure to express the similarity of two voice recordings. These similarity measures are normally estimates of likelihood ratios, attaching weight to two competing hypotheses: same-speaker hypothesis and different-speaker hypothesis. Based on these special properties of automatic systems’ results, the quality of these results can be measured in an evaluation based on a speech corpus. While there exist many evaluations of automatic speaker recognition systems like the NIST speaker recognition evaluations (Przybocki, Martin and Le 2007), the quality estimates based on these evaluations cannot be used to estimate the performance of the systems studied under forensic conditions. The reason is that the results (i.e. likelihood ratio estimates) of automatic voice comparison systems do not only depend on the speakers’ voice features, but also on external factors such as recording/transmission channel, utterance duration and speaking style (Becker, Jessen, Alsbach, Broß and Meier 2010c; van Leeuwen and Bouten 2004; van Leeuwen, Martin, Alvin, Przybocki and Bouten 2006). Any mismatch between recordings analysed by an automatic system has an impact on the likelihood ratio estimate and usually results in an increase in the number of errors made by the system. To avoid unreliable/misleading system outputs, the evaluation of automatic forensic voice comparison systems under realistic conditions is necessary in order to obtain an adequate system performance estimate. In this thesis, the evaluation of systems is based on the GFS1.0-corpus by the German Bundeskriminalamt, which contains forensic telephone recordings of male German speakers.

The author developed two automatic systems based on two different speaker features: cepstral coefficients and formants. While cepstral coefficients are commonly used for automatic speech and speaker recognition systems, formants are currently discussed in the forensic phonetic community as speaker features. The author developed the cepstrum-based system VoCS (Voice Comparison System) and the FGMM System (Formant Gaussian Mixture Model System) in order to provide two different approaches originating from the engineering and phonetic communities. Both systems are based on the basic principle of Gaussian Mixture Models (Reynolds 1992). By comparing these two systems with each other and the two forensic voice comparison systems SPES (Becker et al. 2010a, 2010b, 2010c) and BATVOX (Agnitio 2011)
(which are also based on cepstral coefficients), the following results were obtained:

1. All systems under investigation are able to provide useful information.
2. While VoCS, SPES and BATVOX already provide good estimates of likelihood ratios, FGMM needs further calibration in order to provide such results. This means that the cepstrum-based systems show a superior performance compared to the formant system because they are able to provide more useful information than the latter.
3. Even though the author’s own development VoCS is less complex than the two established systems SPES and BATVOX, VoCS’s performance is the same.
4. All cepstrum-based systems (VoCS, SPES and BATVOX) seem to rely on the same speaker-specific information since a fusion (i.e. a combination) of the three systems does not result in a performance increase.
5. Also, the FGMM System seems not to rely on different information compared to the cepstrum-based systems since the fusion including the cepstrum-based systems and the FGMM System does not result in a performance increase.
6. All four systems under investigation have in common that they are based on short-term acoustic representations of the human vocal tract. While the cepstrum-based systems are able to extract more speaker-specific features, the formant approach is based on the acoustic theory of speech production and is hence easier to understand and to interpret.

All of the automatic systems under investigation do not account for so-called high-level features such as speech behaviour, dialect, foreign accent, articulation rate, melody and linguistic features. Based on this finding, the author’s elaboration of the basic principles of automatic voice comparison systems, and a description of a framework in which these systems can be used under forensic conditions, this thesis can be used as a compendium for automatic forensic voice comparison systems with focus on the special conditions and requirements found in forensics.

References
Becker, Timo, Jessen, Michael, Alsbach, Sebastian, Broß, Franz and Meier, Torsten (2010a)
Automatic forensic voice comparison using recording adapted background models.


The Language of Perjury Cases

Roger Shuy (2011)
Oxford University Press 232 pp

Reviewed by G. Adam Ruther

No man can serve two masters – or so the common wisdom tells us. We are also told that to effectively communicate in speech or in writing we must know our audience. When that audience happens to be a group of educated and experienced experts in our chosen topic, one might think: all the better; write for them, on their level, in their language. But what if a scholar has something to communicate to his or her peers that will be of interest to them but that will also be of great use to another audience – an audience not of scholars but practitioners? Can an author serve these two masters? If the evidence is the work of Roger Shuy, the answer is yes, and Shuy has done it again in his latest monograph, The Language of Perjury Cases.

In this book, the author takes his two favourite audiences on another of his linguistic safaris, this time through federal and state perjury cases. Throughout his scholarship in forensic linguistics, one of Shuy’s consistent themes is that there is a great deal that linguists and lawyers can learn from each other. It is impossible to read this book, however, and escape the conclusion that when lawyers and linguists work together, the lawyers are almost always more enriched by the linguists than the other way around. In this latest, Shuy walks lawyers and linguists, side-by-side, through nearly a dozen perjury cases, on
most of which he consulted as an expert. In using these intriguing case studies, Shuy's book takes on a refreshingly narrative tone, making the book very accessible to his lawyer and lay audiences.

What Shuy has produced is both a scholarly exploration for linguists and – perhaps even more valuable – a linguistic how-to guide for attorneys, investigators and jurists. He very carefully and artfully walks the line between oversimplification and abstruseness that must be observed in order to interest and reach both audiences. For this accomplishment alone he is to be commended. But, of course, this is not the first time Shuy has played to these two crowds at once. *The Language of Perjury Cases* is Shuy's ninth book of this kind in the field of forensic linguistics; he has shown a commitment to lending his linguistic knowledge and talent to those who spend their careers arguing profusely over the meaning of other people's words. These poor souls, whom Shuy has devoted a great deal of his career to helping, are known as lawyers.

I can refer to lawyers in such a way because I am one. What is more, I happen to be the specific species of lawyer with whose linguistics skills Shuy finds the most fault in his book: a prosecutor. In my career, I have learned – and continue to learn through painful trial and error (if you will forgive the pun) – what enormous importance the interpretation of language has in the courtroom. It bears noting as I review Shuy's work that, while I am lawyer, and therefore to some degree a student of language, I am not a trained linguist. In a real sense, then, I am exactly the audience that Shuy is writing for.

Always in search of ways to improve the way I use and interpret the words that are the tools of my trade, I find that Shuy's indictment of most of the other lawyers and judges in his book for their lack of linguistic acumen is not without justification. He depicts many of the lawyers on both sides of perjury cases groping around in the proverbial dark, rarely asking the right questions or recognising the dysfunction in their communication with witnesses and defendants. These problems are certainly not uncommon in litigation, but they are particularly common and salient in the context of perjury and other similar language crimes. Such cases routinely turn on the meaning of just a few words. As the case studies show, a few poorly phrased questions or the mistake of overlooking even a seemingly tiny piece of contextual information in a witness's answer can determine the outcome of a case. The unforgiving nature of these cases may be the reason Shuy finds them so interesting.

The roots of lawyers' deficiencies, however, are not without their own explanation. Put in the simplest terms, they don't teach linguistics in law school. Contract interpretation, yes; statutory interpretation, yes; rules of evidence, yes; direct and cross-examination, yes (if you ask for it); but linguistics? Actual, scientific analysis of language? Sorry, it's not on the menu at most, even
acclaimed, law schools – at least in the United States. If Shuy's work, including this latest offering, shows us anything, it is that basic linguistic analysis should be a part of every good lawyer's repertoire. And when a working knowledge of the basics proves insufficient, legal professionals should be sufficiently well versed to recognise that consultation with a linguistics expert is well advised.

This brings me to perhaps the most valuable aspect of Professor Shuy's work. It is not his actual analysis of the cases he describes, or the critique he offers of the decisions made by the parties, lawyers, judges and jurors in those cases. All of these are easily understood by lawyers, laypeople and linguists alike. He explains, in complete yet comprehensible detail, how basic linguistic tools such as schema theory, speech event and speech act analysis, ambiguity identification, topic and agenda analysis, contextual meaning, concord rules, language and second-culture interference, and the cooperative principle can and should be part of a lawyer's work, not only in perjury cases but in all sorts of legal disputes involving language. *The Language of Perjury Cases* is a case-by-case demonstration of these tools, showing their usefulness in defending and prosecuting perjury charges.

Much like his last book, *The Language of Defamation Cases*, Shuy begins this latest work by setting the legal and linguistic stage through a brief primer on his two fields of inquiry: linguistic analysis and the law of perjury. In his first chapter, Shuy lays out the elements of perjury and explains the most important legal precedents that govern the crime. In his second chapter, he gives a basic overview of how a forensic linguist approaches and analyses language within that legal framework. The chapters that follow are grouped by such themes as ‘Perjury Charges with Inadequate Intelligence Analysis’, ‘Perjury Charges across Languages and Cultures’, and ‘Perjury Cases with Adequate Intelligence Analysis’. In each section, Shuy tells the story of individual cases, conducting a linguistic post-mortem on each. In so doing, he demonstrates when and how a thorough linguistic analysis contributed to the administration of justice in a perjury charge or would have prevented a miscarriage thereof had such an analysis been employed sooner or by a different party.

A professional linguist may find Shuy's analysis of these cases somewhat elementary. After all, it was written so that lawyers could understand it! Linguists' interest in this book will more likely derive from Shuy's description of his boots-on-the-ground application of their field of study. The case studies not only illustrate how linguistic analysis can be applied to perjury cases, but how linguists and lawyers can effectively collaborate and put that analysis to work in court. He encourages members of both professions to accept each other as collaborators and colleagues. Shuy's consistent message is a wake-up call to the legal world and almost a call to arms to the forensic linguistics academy to help
their colleagues in the courtroom whenever the opportunity presents itself. Shuy explains the roots of this idea thus:

Although it may seem obvious to the field of law that sociology, anthropology, economics, psychology, and history can play a role in helping with legal issues, linguistics has had a harder time breaking into the process. Part of this problem may be the fault of the linguists, who have simply not realized how fruitful the area of law can be for applications by their field. But another and perhaps more serious part of the problem is that lawyers and judges use language extremely well, making it less than obvious to them that linguistics has something useful to contribute. Many lawyers and judges also analyze language very well, contributing to their sense of the self-confidence in their own linguistic abilities. But there is also an important difference, whether or not they realize it. Linguists have analytical tools and training that lawyers, even the best ones, do not usually possess. Linguists see patterns and features that are unseen by non-linguists… (157).

My only concern in reading Professor Shuy’s book is that, while it may gain broad readership in the linguistic academy, it may not be read by as many lawyers as would benefit from it. I intend to keep this book as a part of my office reference library and recommend it and lend it out often when colleagues have need of a professional’s perspective on the constant interplay between language and the law. To the linguists who make up my audience in writing this review, I make the same recommendation. The book is an excellent way to introduce even sceptical attorneys to the benefits of consultation and collaboration with forensic linguists on many cases. As Professor Shuy insists, there are indeed great benefits to be realised through the development of a collaborative working relationship between lawyer and linguist.
Analysing Police Interviews: Laughter, Confession and the Tape
Elisabeth Carter (2010)
Continuum 204 pp

Reviewed by Edward Reynolds

In Analysing Police Interviews: Laughter, Confession and the Tape, Elizabeth Carter presents a conversation analytic description of several important structural features of the police-suspect interview. Following the work of Heydon (2005), Edwards (2008) and Stokoe (2009a, 2009b), Carter adds to a growing body of work in conversation analysis (CA) examining the discursive practices of police interviewing. She illustrates the utility of CA with her detailed analyses of police and suspect talk in the interview, the results of which can be easily applied by practitioners. Carter uses this analysis to highlight the way in which the UK’s Police and Criminal Evidence Act 1984 (the PACE Act) is enacted, oriented to, and at times side-stepped throughout the police interview data she presents. The book uncovers elements of the world of the police interview, with real implications for legal practitioners and police. Readers will find the results of this applied investigation a relevant and useful addition to the literature on policing practices.

As a thesis-based monograph, Analysing Police Interviews begins with relevant topical and methodological background prior to presenting the three research themes: laughter, confession and the tape. Chapters 1 and 2 introduce the work,
providing a brief outline of the book along with a discussion of the relevant literature on interviewing. Carter’s overview calls attention to the gap between research and professional practice – a problem which she argues the descriptive work of CA may begin to remedy. She then outlines the approach and research paradigm of CA in Chapter 3, where she highlights the very different assumptions and approaches to data that are used by researchers in CA as compared to the conventional literature on interviewing. In the confusingly structured Part One, including Chapters 4 and 5 (but not, for some reason, Chapters 2 or 3), Carter introduces her first research findings: the analysis of the use of laughter by officers and suspects. She divides the analytic gaze by chapter, focusing on suspects’ use of laughter in Chapter 4 and that of officers in Chapter 5. Carter focuses again on officers in Chapters 6 and 7 and their orientation to ‘the tape’ through ‘daft questions’ and silence. In Chapters 8 and 9, she illustrates how knowledge claims and the minimisation of consequences are used by officers in the elicitation of confessions. Carter’s conclusion in Chapter 10 rounds out the study by providing recommendations for improving practice and adding links to similar work by researchers in phenomenology, ethnomethodology and CA. Carter’s work provides an engaging counterpoint to Heydon (2005) in the emerging literature of CA on the police interview by focusing more on the interactive work of officers.

In the context of policing studies, the details of actual interviews are often overlooked in the clamour to improve interrogation methods (Heydon 2005). However, it amounts to a guessing game to try to improve interrogation methods if there are no adequate empirical studies of real interrogations; this is where Carter’s research fits in. After a brief overview of the book in Chapter 1, Chapter 2 sets out the literature on interviewing in the police context. In this chapter Carter provides a good description of the legal context enshrined in the PACE Act, an overview of sociology and criminology work on police culture and methods, and discussion of research on police training in interviewing technique. She makes the argument that confession is treated as the goal of the interview, despite the rarity and difficulty in obtaining such confessions (p. 13). This chapter sets the context for a comparison between the rhetoric and the reality of police interrogation. However, for methodological and practical reasons it does not provide a detailed and up-to-date discussion of interview technique (a reasonable omission in the wake of recent legislative changes in the UK), saving the description of interview techniques for the data itself.

Chapter 3 sets out CA’s basic assumptions, the data collection strategy used in this investigation, and a brief introduction to some basic techniques of CA. Carter describes CA’s distinctive focus on detailed, naturalistic investigation of the way in which language (‘talk’ in CA jargon) is employed to accomplish
action. She provides a partial history of the discipline and briefly discusses CA's inductive analytic approach in the context of data collection and the selection of the phenomena to be analysed. (Readers unfamiliar with CA may not find that this rather dense, jargon-laden chapter sets them up particularly well with what they need to read the following chapters; I would recommend instead Sidnell 2010.)

The research findings begin in Part One of the book, including Chapters 4 and 5, which analyse the organisation of laughter in the police interview. Carter highlights the way in which the asymmetry built into the police interview partitions the uses of laughter differentially for suspects and officers. In Chapter 4, the analysis identifies how suspects employ laughter in order to minimise the potential for conflict inherent in challenging statements. She illustrates suspects' use of laughter as a resource to either support their position or challenge the officer's position, using laughter to downplay the conflict while at the same time attempting to pull off a challenge. Chapter 5 explores the use of laughter by officers to both acknowledge the interactional tension between the demands of co-present talk and the requirements of PACE or, more worryingly, at times side-step it. She highlights several cases in which officers employ laughter to step outside the framework of the interview and in doing so breach certain provisions of the Act. This particular finding reproduces research in the British context by Ainsworth (2008), whom she overlooks. With her illustration of the methods used by officers to downplay their breaches of PACE, the analysis further underlines the relevance and utility of CA in applied studies of policing and the law; see also Travers and Manso (1997). Carter's findings bring the sociological investigation of policing closer to practitioners by providing specific, concrete examples and recommendations for future practice and research.

Part Two of the book explores the officer's orientation to the suspect and to the 'silent participants' – juries and judges that may later hear the recorded interview. In this more linguistically focused part, analysis begins in Chapter 6 with an exploration of the way in which the requirements of enacting a police interview are sometimes at odds with everyday rules of interaction. Carter points up the 'semantically redundant' organisation of much of the interview structure required by PACE and the trouble this creates for officers. Echoing Stokoe and Edwards' (2008) analysis of 'silly questions', she highlights the way in which parts of the interview can be treated as 'daft', obviously conducted simply 'for the benefit of the tape' at the cost of a 'smooth-flowing' interaction with suspects. She argues that interview requirements and training ought to be amended in order to better align suspects' rights and adherence to PACE, making them converging rather than competing demands on the officer. In Chapter 7 Carter explores how officers use 'uninitiated third turns' for the
benefit of the silent participant in order to treat some part of the answer as the ‘official’ version. She teases apart the different sorts of third turns used in her data – both everyday conversational repeats and ‘institutional’ repeats. Part Two illustrates the practical accomplishment of a taken-for-granted feature of police interviews: their nature as something designed for later listening as a part of the legal process. From a CA perspective, the analysis in this chapter is somewhat pedestrian and out of date – e.g., the notion of ‘redundancy’ implying some parts of talk might not have some relevance in contrast to recent CA presumptions (Schegloff 2007). Setting aside the semantic quibbles, Carter’s practical analysis in this section can be seen to identify specific ways in which police interview training can be improved in order to remove redundancy while retaining PACE compliance in order to benefit the process of the interview.

Part Three explores the much-glamorised confession in the police interview. Carter’s data supports the often overlooked point that confessions in police interviews are exceedingly rare: less than 5% of her examples include any form of confession (n = 150, taken over three years). She juxtaposes this observation with research on the rhetoric of police work which emphasises officers’ orientation to the interview as sources of confession, arguing that the large focus in the literature on improving rates of confession creates an environment that puts the rights of the suspect at risk. With this cautionary note in mind, her analysis in Chapter 8 explores the way in which officers employ knowledge claims in order to challenge the suspect’s veracity – challenges that result in ‘confessionary turns’ (i.e. partial admissions or full confessions). Carter highlights officers’ use of outright statements of knowledge and the more subtle presuppositions that can be built into questions in order to elicit a response from uncooperative suspects. Chapter 9 explores another technique that officers employ in their attempts to obtain confessions: minimisation of the seriousness of the crime. In this chapter, Carter describes how officers can downplay the potential consequences of a crime as a method to obtain confessions. This analysis was a straightforward demonstration of a PACE-compliant interview technique taught to officers as a part of their training. Readers should note that in line with CA principles, Carter does not – and cannot – provide recommendations as to the ‘best’ methods for obtaining such confessions; her analysis is instead exploratory and descriptive. Research questions about interview technique are better served by predictive experimental work. Nevertheless, as in the case of Chapter 8, some analyses indirectly suggest additional methods for an interviewer’s toolkit and further avenues for research.

Carter closes the book in Chapter 10 with conclusions and discussion. Here she brings the findings of the analytic sections to bear on recommendations for
training and research that might both assist police work and enhance compliance with the spirit of PACE. Based on her results, Carter recommends that police interrogation training methods be founded on the results of real, empirical research. As an example of this, she spotlights the way the police use of laughter concords with more conversational interrogation methods. Further in this chapter, she discusses recent developments with the alternative PEACE interview model and adds some discussion of how her findings accord with other ethnomethodological research on police interviewing and interrogation. I was at times confused by the references to symbolic interactionism: Carter appears to want to juxtapose symbolic interactionism with CA without giving an introduction to the former. I was also disappointed that it took until Chapter 10 to link the differential laughter rights detailed in Chapter 2 to ‘category predicates’, the way different actions are tied to categories – e.g., officers can laugh to break the interview frame, suspects can laugh to challenge the officer, etc. It might have better served the analysis for predicates to be included in Chapter 2 in the discussion of ‘role’ or left out altogether.

From the perspective of a conversation analyst, Carter’s book left me wanting more, in both good ways and bad. The description of the practices of enacting the PACE Act, orienting to the silent participant, and methods for eliciting confession represented fairly straightforward sociolinguistic description – conversation analytic ‘stamp collecting’. Carter’s analysis lacks a coherent description of the local social order, instead presenting a grab-bag of practices, all drawn from the interview but without a consistent analytic thread. Further, the analysis falls prey to ‘sequentialism’ (Heritage 2012), as it focuses on a very narrow domain of CA, thereby making the potential for the description of the enactment of the PACE Act in and as the police interrogation falls short. The upshot of the findings might be most relevant to practitioners, police and lawyers in crafting interview practices better aligned with the competing demands of legislation, policing and citizens’ rights.

*Analysing Police Interviews* provides a valuable application of CA to an underexplored research domain. This work serves to point out how little we know of the police interview and how much more there is to learn. Carter demonstrates the applied utility of CA in the way she employs real data from real subjects in a high-stakes setting to illustrate some of the unique features of this part of the legal system. As a part of the growing number of ‘applied’ CA studies – CA research is *ipso-facto* ‘applied’ – it is CA’s findings that need application. This work takes up Vrij’s (2000) call for researchers to engage in more naturalistic inquiry in the area of interviewing. As it is based on real data, the scenes in the analysis should resonate with readers with police or legal backgrounds, and it is precisely this sort of evidence-based approach to making
practical recommendations for police practice that gives this work its strength. It should be immediately clear that officers sidestepping PACE in the course of interviewing is both a worrying trend that needs correction and an avenue for possible judicial appeal. With such clear practical relevance, lawyers, linguists and psychologists alike will find this a useful application of CA to policing studies.

References


