



COMPARING VOICES: SPEAKER IDENTIFICATION WITNESS SEMINAR

Transcript of a Witness Seminar on Speaker Identification, Maastricht University, the Netherlands (in person, 17 June, and online 9 November 2022).

Anna Kvicalova & Karin Bijsterveld
(Eds.)



The editor(s) 2023. This is an open access publication, licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this book or parts of it.

Graphic design by Tomáš Janků (Orbis litera). The cover image design was created in the AI programme Midjourney (2023).

PLEASE CITE AS:

Kvicalova, A. & Bijsterveld, K. (Eds.) (2023). *Comparing Voices: Speaker Identification Witness Seminar*. Maastricht University.

THIS PUBLICATION HAS BEEN GENEROUSLY SUPPORTED BY:

- Research Stimulation Fund, Research Institute Faculty of Arts and Social Sciences, Maastricht University (RSF Grant 2022, Round 1)
- Foundation Scientific Research Limburg (SWOL Event Grant 22.003)
 - Centre for Theoretical Study, Charles University, Prague
- Czech Science Foundation (research project The Second Sense: Sound, Hearing and Nature in Czech Modernity 20–30516).

ISBN:

978-90-9037363-8

TABLE OF CONTENTS

01	Introduction	4
<hr/>		
02	Transcript of the witness seminar	9
<hr/>		
03	Transcript of the interview with Peter French	45
<hr/>		
04	Biographical notes on the witness seminar participants and editors	71
<hr/>		
05	References	74

01

INTRODUCTION

INTRODUCTION

The event

On 17 June 2022, a small group of specialists in forensic phonetics and acoustics gathered at Maastricht University in the Netherlands for a witness seminar on the history of speaker identification. The group included Angelika Braun, Herbert Masthoff and Maartje Schreuder. A fourth invited expert, Peter French, cancelled his on-site participation shortly before the event due to travel restrictions caused by an airport strike in the Netherlands, while a fifth invited expert, Ton Broeders, was unable to attend for personal reasons. The organizers suggested an alternative course of action which was agreed upon by the other participants: Peter French was given access to the witness seminar transcript and then interviewed separately in November in an online conversation with the organizers. The interview with Peter French constitutes the second part of this document and represents an important addition to the witness seminar as it responds to the on-site discussions in Maastricht.

The event was co-organized by Maastricht University (the Science, Society, and Technology Studies research programme) and Charles University in Prague (Centre for Theoretical Study) and chaired by historians of science Karin Bijsterveld and Anna Kvicalova. To ensure an open discussion among the participants, the seminar had no audience.

The witness seminar was held in the morning and was followed by a workshop in the afternoon. During the workshop, five humanities scholars presented research papers on the history of forensic sound analysis in Europe and the USA.¹ Although the participants in both events met together for lunch, they did not attend each other's meetings. This transcript represents the first opportunity for historians to access the memories and reflections of a selected group of speech scientists shaping the field of forensic phonetics and acoustics over the past decades.

¹The papers written for the workshop are being published in Bijsterveld K. & Kvicalova A. (Eds.) (2024). *Forensic Voices: Cultures of Sonic Identification and Detection in the West*. Special issue of *Sound Studies*, 9(2): forthcoming.

The method of the witness seminar

The witness seminar is a moderated discussion in which a selection of individuals with expertise in a particular area is invited to exchange memories of and experiences concerning the history of methods and practices in that field. When used as a research method in the history of medicine, science, and technology, its objective is to juxtapose different perspectives on specific subjects, and thus both enrich and challenge public histories. The purpose of the witness seminar is not to create an authoritative narrative of the history of the field, but to uncover untapped sources and open up unexpected avenues to further research. It adds personal perspectives to well-known events, points to important names and connections, and offers unexpected interpretations beyond the scope of published sources.² The witness seminar never simply records history, but it reveals a multiplicity of *histories* of the field as these are constructed in individual memories, opinions, and interactions.

Comparing voices

The witness seminar had two objectives. The first was to learn more about West European speaker identification methods by drawing on the participants' personal experiences in the field. The second was to gain some new insights in speaker identification practices in Eastern Europe around 1989. The year 1989 marks the end of the Cold War, a time period during which speaker identification methods were explored by police *and* secret intelligence agencies in the countries of the Eastern Bloc,³ and professional contacts across the Iron Curtain were very limited. Because several of the participants had worked on speaker identification in the late 1980s, and Western and Eastern European experts had more opportunities to visit each other in the years immediately following 1989, we hoped to acquire a better understanding of the commonalities and differences in speaker identification practices on both "sides" of Europe.

We chose experts from (the former West) Germany, the Netherlands, and the United Kingdom, who represented different approaches and national traditions of doing forensic speaker identification, but all of whom have been involved in decades-long professional debates within the field of forensic phonetics and acoustics both within and beyond the International Association of Forensic Phonetics and Acoustics. In addition to our historical focus on the immediate post-Cold War period, we were interested to identify and discuss the key methodological, legal, and ethical issues in the development of the field. We decided to invite a limited number of experts to provide ample time for exploring the relevant histories and voices in those histories in-depth, to which the additional interview with Peter French further contributed.

² See, for example, Maas H. (2018). The Method of the Witness Seminar. *History of Political Economy*, 50(3), 571-577; Tansey E. M. (2006). Witnessing the Witnesses: Potentials and Pitfalls of the Witness Seminar in the History of Twentieth-century Medicine. In: Doel R. E. and Söderqvist T. (Eds.) *The Historiography of Contemporary Science, Technology, and Medicine: Writing Recent Science*. Oxon: Routledge, 260-278; and Nicholls E. J. (2020). The Witness Seminar: A Research Note. *Qualitative Research*, 00(0), 1-8.

³ Bijsterveld K. (2021). Slicing Sound: Speaker Identification and Sonic Skills at the Stasi, 1966-1989. *Isis*, 112 (2), 215-241; Kvicalova A. (2023). Sound on the Quiet: Speaker Identification and Auditory Objectivity in Czechoslovak Fonoscopy, 1975. *Technology and Culture*, 64(2), 379-406.

Editing and publishing the transcript

We made an audio but not video recording of the witness seminar. The transcript of the oral interactions went through several steps. As soon as a draft transcript was ready, we shared it with all seminar participants. Each of them got the chance to correct the text and add references if applicable. At this stage, the participants were granted the right to edit any of their answers should they consider them unsuitable for publication. None of the experts deleted any of their answers, though – they mainly language-edited their answers. The edited transcript was then shared with the fourth participant, Peter French, who was, as said, interviewed separately, but had the option to respond – directly or indirectly – to the answers of the other participants. Similar to the other participants, he was allowed to edit the draft transcript of his interview. Granting the participants the right to approve and edit the transcripts before publication enabled an open discussion. In the final stage, brief explanatory footnotes and references were added to both parts of the transcript. The present version of the transcript stays as faithful to the original interactions as possible: it was only lightly edited, and it preserves the oral nature of the seminar. The original sound recording of the conversations will not be made public.

The open-access publication of the witness seminar provides a unique source of information for academics as well as speech science practitioners, as it offers a glimpse into the field's development as perceived by experts with different national backgrounds and professional experiences. As such, the transcript contributes to the corpus of oral history sources on the history of 20th century science and technology.

Structure of the seminar

The witness seminar was structured around five main questions, which were formulated by the organizers and shared with the participants a few weeks before the event:

1. What is the most memorable case you worked on and how exactly was your expertise involved?
2. What did your work routine look like *in practice* (in terms of technologies, procedures, cooperation, workspace etc.)?
3. Which event(s) do you consider crucial for how your field developed?
4. How did opinions about the best and most authoritative way to do speaker identification change over time?
5. To what extent did speaker identification experts in the West have contacts with speaker identification experts in Eastern European countries?

Despite the structured nature of the seminar, the participants had the opportunity to react to each other and raise issues which they considered relevant for the topics that were being discussed. Additionally, the participants were encouraged to bring with them documents and objects they found particularly relevant for the topic of the seminar. Some of them made use of the opportunity and brought published, although not widely available, titles on forensic phonetics and acoustics.

We are deeply grateful for the time and efforts the participants Angelika Braun, Peter French, Herbert Masthoff and Maartje Schreuder invested in this witness seminar on speaker identification. We would also like to thank the Research Institute of the Faculty of Arts and Social Science at Maastricht University, the Foundation Scientific Research Limburg (SWOL), the Centre for Theoretical Study, Charles University, Prague, and the Czech Science Foundation for generously funding the witness seminar event and publication.

Note on the publication of personal data

All the participants in the witness seminar signed an informed consent form in which they agreed to the publication of this transcript for research purposes. Where they mentioned third persons by name, we have made these names public only in the cases of scholars and researchers whose names are already known from their own academic publications. These concern German, English, American, Polish, and Russian scholars. Their names are mentioned strictly in connection to their published work and no personal information is disclosed. The names of former members of the German police and persons who have not published as scholars in the field are anonymized. The information about court cases is anonymized except for the names of world-famous cases (e.g., the Rodney King case, Zimmerman vs. Florida, the International Criminal Tribunal for the former Yugoslavia), which received wide media attention.

Anna Kvicalova & Karin Bijsterveld

02

TRANSCRIPT
OF THE WITNESS
SEMINAR

COMPARING VOICES:

SPEAKER IDENTIFICATION WITNESS SEMINAR

The transcript of a Witness Seminar held at Maastricht University,
the Netherlands (17 June 2022)

PARTICIPANTS

Chairs:

Prof. Dr. Karin Bijsterveld

Dr. Anna Kvicalova

Participants (17 June 2022):

Prof. Dr. Angelika Braun

Dr. Herbert Masthoff

Dr. Maartje Schreuder

Apologies:

Prof. Dr. Ton Broeders

Prof. Dr. Peter French

Bijsterveld:

Another welcome. Now the official one. I am really *really* glad that you are here and appreciate that you took the time to travel to us despite all the hurdles nowadays. Thank you very much for being here.

We make recording as you know, and we also make notes. I'm the formal chair, but Anna will simply jump in if she has any questions or if I forget something, to also keep an eye on time. I would like to explain once again what a witness seminar is. [...] This is really about a moderated conversation about a particular field. Sometimes it focuses on one event, sometimes on an episode, so we focus on quite a long episode and for us it's also very important to hear about your contacts with the East German and, actually, the Eastern part of Europe, as far as you had them, because I know that is different for different people here in in the room.

It is not about establishing the "right" version of history, but about hearing different opinions, different views, different experiences, different moments, and as we explained before, we try to align that with other sources as well. The witness seminar is also seen as a style of research, because it's an easy way to hear about your experiences.

We have five questions, so we have about half an hour for each question. We also have a break - if you would like to have a break earlier that's possible as well. In addition to those five questions, we have this sixth question about whether you would like to share something with us. Maartje [Chreuder] already sent materials, but we would like to reserve some space just to see whether you would like to share something with us that you think is important to understand the field. So that's basically what I would like to share. Anna, is there anything you would like to add?

Kvicalova:

Perhaps we should mention (or we can also discuss that in the end) that Peter French wasn't able to come, but he is still, at least partly, part of this workshop, because we would like to interview him later based on the transcripts of recordings from the meeting today. But we will discuss the details later this morning. I only wanted to mention that he cannot be here, but we would like to get him involved after the workshop (...).

Braun:

Excuse me, do we get to see his comments as well?

Kvicalova:

Yes, of course.

Bijsterveld:

Yes, we explicitly talked about that yesterday. We will first make a transcript of your discussion, then you will have the chance to correct it, edit it, add something if you would like to. And then, only then, we share it with Peter. Then we have a talk with him, online, record it, make a transcript of that in its entirety and then we share it again with you. That comes most close to how the situation would have been would he have been in the room. Still, we decided not to do the workshop in a hybrid format, because that really works differently than doing it in this way. At least that is how we think about it and why we would like to do it in this way. We also have asked Ton [Broeders] to do it in the same way, he hasn't responded yet. [...]

Masthoff:

I think that's fair.

Bijsterveld:

We would like to start with the most important case or the most memorable case you worked on. Of course, we understand that you cannot share everything, but perhaps you can also say something about how confidential your work is, or remains over time, and, how you use your expertise in that particular case. I don't know whom of you would like to start. Perhaps Angelika [Braun] – we can start with the most senior one.

Braun:

I'm afraid I cannot answer this question in a simple way by citing just one case, because there are many memorable cases in different respects. For instance, there was one case, where later on the investigators told me that my report had most likely saved the kidnapped person's life. So that, of course, is memorable in the sense that I will always remember it.

That was a kidnap case where they had a suspect, and the suspect had been followed and it was quite certain that for days the suspect had not been taking care of the hostage. And so, they needed a report to apprehend the suspect and question the suspect and when they did question this suspect intensely, the suspect revealed the whereabouts of the kidnapped person. And since that person had not been taken care of for a number of days, it would not have gone on for very long. That is the reason why they sought the report and then they had the report to get an arrest warrant. That's how that probably saved a person's life.

Then, of course, there have been a number of cases and one in particular with a fairly high international impact. That was in my case The Rodney King case. Rodney King, being the black person, who was beaten up by the white Los Angeles police in, I think, 1987 or so. Then I was not part of the criminal trial, I was part of the civil trial, where I think for the first time the audio of the video in the Rodney King case, was shown around world. The audio did not receive much attention until the civil trial. The reason that I got involved was that I met another expert who was involved in that trial at a conference in Düsseldorf and I gave a talk on what forensic phoneticians do and, later on, they walked up to me and said: "Well, would you like to have a go at the Rodney King audio, because, one thing we need to know is when, at what time, the Taser was being used." They were wondering whether the Taser noise could be identified from that audio. As it turned out, that could very well be done. Plus, they wanted me to find out, or establish, whether the word *nigger* had been used at any stage during that event. That is how I got involved in that one and that was the most memorable one with respect to international impact I would say.

The most emotionally challenging one, and the most memorable one from that sad respect, was having to work on a cockpit voice recorder, when several minutes passed between a head-on collision of two airplanes and the actual end. So we had a couple of minutes of distress. And that, I would say, was the most emotionally challenging case.

Probably the most interesting from a professional point of view is a very recent one, where there were already three reports from three different experts and mine was the fourth. That was of course, an extremely interesting case from a professional point of view and so it is memorable as well. I'm sorry that I do not have a brief answer to your question.

Bijsterveld:

No, this is perfect because this is exactly what makes it rich. This answer shows, indeed, how cases can be memorable for different reasons. We may have some in between questions.

I have one very specific question about the crash: What was the issue there that you had to figure out?

Braun:

The wording, because it was distress speech and that, of course, is nothing like normal speech. And also background noises.

Bijsterveld:

...which could say something about the cause?

Braun:

The cause was quite clear, it was an immediate head-on collision of two planes.

Kvicalova:

Were you also ask here, or in similar cases, to assess the emotional state of the pilot? As far as my historical knowledge is concerned, this was routinely done, for example, in Czechoslovakia when the experts were called to do cases of plane crashes and assess the emotional state to determine whether perhaps this was not a plan, whether the stress was authentic. This was not a part of your job?

Braun:

No, because we usually, in accordance with the Code of Conduct of the International Association for Forensic Phonetics and Acoustics, do not do this. Personality assessment, and, in my view, statements about the emotional state come under that heading. What we do, is describe what's going on. For instance, if the speech behaviour changes, we will describe that change but we will not generally jump to conclusions like "this is stress". We will describe the change and the attribution is not really something that is in the centre of our work.

Bijsterveld:

Thank you. Did you say what was the year of the accident?

Braun:

It was in 1997.

Bijsterveld:

What strikes me about the Rodney King case is also that you focused on the Taser noise. So it is not only the voice you focus on.

Braun:

It is also background noises and, very often, a fact-finding mission, because,

for instance, if a victim manages to call 911, and then is suffering further attack, then that attack will be captured on the 911 recording. Like for instance in the Trayvon Martin case, where background noises played a very important role.⁴ Peter will probably have to say something about that. In those cases, our work is also to do fact-finding. Like for instance: this person must be approaching the microphone because the voice is getting louder at this minute. Sometimes you can hear the knife going through the fabric if there is a stabbing. Oftentimes, it's the question of who shot first and there is a sequence of shots being fired, so we also do fact-finding.

Bijsterveld :

In case of the Taser noise, did you also do comparison with existing recordings of that sound?

Braun:

Yes, I asked for a reference recording in order to determine whether it was at all possible for the Taser to be recorded by that camera. They actually shut off a small part of Los Angeles and had a helicopter hover over the scene because there was also the helicopter noise and you had to determine whether the Taser could be made visible or audible despite the helicopter noise. So they had a helicopter hovering over the scene and they had somebody with a Taser producing noise for comparison purposes.

Bijsterveld:

Did they also use the same type of camera?

Braun:

Yes, absolutely. If not that individual one, I'm not sure about that. But the same type of camera and type of microphone, the same type of Taser, same everything.

Bijsterveld:

Thank you very much, we may return to your cases, but I would like to distribute the attention. So let us give the floor to Herbert [Masthoff].

Masthoff:

Yes, certainly. I agree with Angelika that picking out a single case is rather difficult. I've been doing these expert witness reports for close to 30 years and every case is, from a phonetic point of view, memorable and interesting because every single case offers unique and particular circumstances.

I have two cases that I'd like to mention. I find them memorable, but, in a way similar to you [Angelika] they are not memorable because of any distinctivity in the voices or so, but for circumstantial reasons.

There was one case I worked on in for the Landgericht Leipzig. There were two defendants, it was about human trafficking, and the job was to compare two sets of telephone intercepts, were one of the sets consisted of some 21,000 individual

⁴ *State of Florida v. Zimmerman* (2012) was a high-profile case in which George Zimmerman, a neighbourhood watch captain shot and killed a 17-year-old African American teenager named Trayvon Martin. In analysing a scream for 'help' captured in the background of a recorded 911 call, two different methods – auditory-acoustic and automatic speaker recognition – were used with varying results.

phone calls. The first meeting I had with the judge was: “Can you do this?” “Sure, it sounds like a routine type of job.” “How soon would you be able to produce your report?” I said, “it’s gonna depend on the amount of material involved so, what are we talking about?” “Well,” he tells me, “It is 21,000 intercepts.” And of course, my response was, we need to bring that number down to have a chance to produce results here in a reasonable time frame and he realized that. And I told him: “Look, you need to bring down the number as far as possible, so we ended up with 120 something calls.” Those were among the high-profile calls, because, on the one hand, he wanted to have as much material in there as possible but, on the other hand, within the time constraints of the court case. So we agreed on about 120 or something calls.

That is about the top of the list as far as numbers of phone calls involved, the 21,000. I have had other cases that had 150 or so calls but to start off with 21,000 that is absolutely unique. What also makes this case memorable, though, is that it came to a point where I presented my report and I produced some diagrams, spectrograms, plots of fundamental frequency analysis and sonograms. I explained “Okay, this diagram shows between time X to time Y the following phenomenon...” The defence argued: “Well, that’s all fine, but, in order for us to understand your report, we request that not only selected portions of the recordings are documented, but that every millisecond will be acoustically analysed and graphically represented.” So, in other words, they were asking diagrams for every single millisecond, which put me in a position that I had to inform the court about general and acoustic phonetics as related to voice analysis, and voice comparison. Essentially, that the diagrams for every single millisecond doesn’t mean anything. Usually it is the broader acoustic context around specific point of time that will exhibit the speaker specific features in a segment.

I had to go there to Leipzig, I think it was four times or five times altogether, to present my report. Following up on the defence application, the court ruled that not every single millisecond will have to be documented, but that the documentation’s purpose should be to substantiate [the findings] in a meaningful way and be presented in a way, that non-experts can relate to. I found this case memorable because this ruling was very helpful for me in cases thereafter.

The second case is more memorable due to the circumstances. That was a case presented to the court of the city in Wuppertal. A gang of tomb raiders were involved here, a gang of five, and their ambition was to steal metals, bronze in particular. The job here was to compare two sets of telephone intercepts, the same as for Leipzig, which was two sets of telephone intercepts. One set known, another set of unknown, so I went ahead and did the voice analysis and the voice comparison. The voices involved were rather distinctive, so it was a rather straightforward job. And the language was Polish, by the way, so the judge proposed to me to include a translator, a native speaker of Polish, which I agreed to, and my results were supported by the findings of the translator. In that he said, “Well these guys are, according to their dialectal features, from the greater Warsaw region, and they have the same level of education, and the grammar is similar here so that matches up with the phonetic side of the analysis.

So that was fine. Our job was terminated, I went to Wuppertal along with the translator, we presented the case, and we were finished at that point. Two days later, I got a phone call from the translator. He said: “You know, I found this very interesting. I went over the conversations of these guys, who were going back and forth between Poland and Germany. So what I did, I infiltrated *Facebook*.” The Polish section of *Facebook* I guess that must have been. “And guess what happened? I found these guys and their conversation on *Facebook*. And not only that, but they had also taken pictures of each other while they were stripping the tombs.” I said: “Look, I think we should tell the judge about this, thank you very much.” So I called the judge and I said” This is what happened, Your honour.” And he said: “Oh, wait a minute, how am I going to introduce this information in a liable way? Hold on, just freeze, don’t talk to anyone about this.” He called me back the next day and he said: “I have a way to do this, this will be the *coup de grâce* for these guys.” We talked about the case a couple of weeks later and he said everything was fine with the case and special thanks especially to the translator. This input had been extremely helpful.

Bjsterveld:

You talked in the first case about making a report meaningful also in a way that non-experts could relate to it. Does that happen a lot? This may be something for the others to step in.

Masthoff:

I have to say that that is one of my top objectives in presenting the report. Our job is, or at least that is how I understand the job, to offer a helping hand to the courtroom, and it makes no sense to present your results... There is a fine line between being scientific here, on one hand, but, at the same time, being able to transmit the scientific findings to others to be able to relate to it. So yes, I try to find that line and I think it is very important because otherwise, I am not doing the right job.

Bjsterveld:

Are you hesitating sometimes about how you present it?

Masthoff:

Yes, I do, yeah.

Bjsterveld:

Can you give an example?

Masthoff:

It’s hard to say. It depends on the individual case and on the individual findings. It’s very hard, for example, to explain to a judge that while your vocal tract may exhibit specific resonances, those resonances sometimes are and sometimes are not displayed in a sonogram or spectrogram. So here you go, a resonance in the vocal tract and the spectrogram and the frequency of 2100 Hertz or so. The point is to break it down to a level that may be understood by a non-expert without leaving the scientific or forgetting about the scientific basis.

Braun:

It has to be understood by the judges and the attorneys that are in court, but also by a fellow expert. That's probably the line that you are describing and if there is a second expert, he or she should also be able to grasp what you found and understand what you were saying, but so should the judges and the attorneys and maybe also the defendant.

Masthoff:

For the colleague to be able to reproduce the findings. Reproduction of the findings for fellow experts but at the same time for the non-expert to be able to understand what all the findings mean, and how to explain them, so they get a picture of what happened.

Bijsterveld:

And is there always a second expert?

Braun:

In Germany hardly ever, because we don't have the adversarial system, but the inquisitorial system, so the expert is appointed by the court. That is supposed to guarantee impartiality.

Kvicalova:

I would like to go back to how you make your evaluation meaningful to non-experts. Are there any general strategies which you adopt and which you think make it easier to communicate the results? Such as speech sonograms: are they helpful, or rather not? How do you actually make the judges and non-experts understand?

Braun:

I follow sort of a dual strategy. I write down one version for experts, where I use terminology, and then I explain that terminology, or try to explain that terminology, in lay terms. For instance, if you are talking about the *glottal stop*. Which is the sound, "uh-oh", "ah", "uh", and so forth. You can call it a *glottal stop* but then, in parentheses or in another sentence, explain that a *glottal stop* is a sound that is made, in such as such a way and give an example and so on. We do both, we use the term for the expert or for anybody who is knowledgeable to understand what exactly it is being said, and an explanation is added for the lay listeners and for lay customers. You just say: "This black bar shouldn't be there, but it is." "This is the significant bit about this illustration." And for the expert, you say you call that black bar a *formant*, and you call it formant number 1, 2, 3, 4, 5, you put it in exact terms.

Masthoff:

I agree. The situation in the courtroom is very comparable to teaching undergraduate students. I found this very synergetic, in a sense, to be a teacher in the classroom, and take that experience and transfer that to the courtroom. And the experience that you pick up in the courtroom and take it back to class. I found that to be very helpful that the strategy that I was able to apply in the classroom would also work in the in the courtroom and vice versa.

Bjsterveld:

Thank you very much. We might return to these issues, but now I would like to give Maartje the chance to say something about her memorable cases.

Schreuder:

My case is a case that is memorable for all the things that went wrong. It was a 2009 case of women trafficking. Normally, before I start a case and during the case, I do not know anything about the case. That is an important aspect of how we do the cases in Maastricht here. Following Ton's [Ton Broeders] strategy, we are completely blind for the background of the case, so we just look at the evidence itself. In that case there was material in Nigerian languages, namely Edo and Pidgin English and we had to record reference material from the suspect because there was no reference material available.

The suspect had just been transferred from England to the Netherlands and we went to Zwolle to record the recordings and before that we had an educated native speaker to translate our scripts into the Edo language and Pidgin English. And we recorded these reference materials through the telephone line. I was there with the suspect and with an intern, I think, and someone else was present at Maastricht University recording via the phone line with an educated native speaker present for me to talk to the suspect through the phone. And we have a script that consists of different parts. [The first part was to] read something out, that was a sort of informed consent. Then we had some semi-spontaneous conversations, and some other tasks.

We had those scripts translated, and we were ready to record. I instructed the suspect: "We have the scripts, and this is what you have to do, we have these tasks we do in Edo and this tasks in Pidgin English." But he said: "I don't speak those languages." I texted my colleague who was in Maastricht and asked the conversation partner to just start talking in those languages and see what happens, so he did. But the suspect said: "I don't understand him." So we decided to just do it in English.

Normally, what we do when we make these recordings is that there is someone present who knows the suspect and especially the voice of the suspect. For instance, that is a police officer, or a warden of the prison, who really knows the person. But this person had just been transferred here so there was no one in the Netherlands who knew him, so no one was present who knew him. Having someone present who knows the person works normally preventative so they will not disguise their voice. But also when they would disguise their voice, the person [police officer or warden] would notice so. It works in two ways, and normally it doesn't happen. We have hardly ever had someone disguise his or her voice, but in this case, we immediately suspected that he did so, because he was talking in a very high pitch, almost feminine voice, with a very posh English accent and I don't know if you know the movie Mrs. Doubtfire, but that is the way he talked. So we had this Nigerian man saying: [imitating the high pitch voice]: "Hello, yes, I'm fine." And we thought: "Hmm." We looked at each other and we just continued with the recordings, but immediately after, we informed someone from the English serious crime squad, who was involved in that case

and who knew him, and we asked him: “Could you please describe his voice and his speech?” And what he told us was completely different from what we heard on that recording. And afterwards we sent some fragments of that recording to that same officer, who had a big laugh and said “No, that is not how he normally sounds”. So we informed the investigative judge, and we said this is not useful material, we cannot do a comparison with that. They told us to continue the analysis, so we analysed the material described, compared it and said, well, we could not draw a conclusion. But it was a very interesting experience for that reason. It was woman trafficking, as I already said.

Bijsterveld:

Thank you very much. You said “normally, we don’t know anything about the case”, and another part of the sentence was “this is how we do it in Maastricht.” Is that special?

Schreuder:

I find it important that the analysis that you do is independent of any other information. Knowing that it is women trafficking could potentially make me look harder for similarities, if I find there is a very important reason to do the case better, while you should always also consider differences. That is of course something you always strive for, but there is the cognitive bias... Angelika laughs, we will have this discussion.

But the problem with cognitive bias is that people are not aware of it, and they tend to think that you can just eliminate that by willpower and introspection, which, unfortunately, is not the case. I’m very certain that I want to prevent myself from being biased by any information, including the information that one part of the material is from the suspect and another part is questioned material, so I don’t know what source is what. I also want to have other speakers included, so that I also know that I can make mistakes in the comparisons, so that there are definitely other speakers involved. So that I have to distinguish and compare.

I always have an assistant who prepares, who shields me from any information that I don’t need and prepares a line-up of fragments of the material. The questioned material, the reference material and some fillers. And that’s a line-up with a random order of fragments, just numbers, no other information, and my task is to compare all the fragments and to group them into different groups. That is the first stage, a blind stage. In the second stage I do get some information. There is another blind stage in which I only get information about what material is the reference material. And that I can see whether that changes my initial grouping or not, and at every stage I record my grouping so that is always recorded, so if I made mistakes that will be in my report. Then there are some more stages and at the end stage, the full information that is relevant is revealed so that I look, I do the full analysis and look whether there could potentially be some information that could change something, for instance, information that there were brothers involved or maybe that there were recordings from the same person at different ages. Or there could be all kinds of things, also differences in phone line or other recording devices. There could be all kinds of information that could potentially explain differences or similarities.

But usually that should not change the conclusion itself, only the strength of the conclusion, if you talk about probability. There is also a fuller range of the material that cannot be used in the blind stage, because in a line-up you cannot use all the materials and not 21,000 fragments, but also not 125, so 20 max. So you cannot include all the material, such as material of bad quality or very short material or different languages or different speaking styles. If there is, for instance, one fragment, where someone whispers, it cannot be usually included in the line-up, only if there is also whispering in other material then you can do it, but otherwise not. It does not prevent every form of bias, but I think it is an important stage. I find that important also because if the police provide information or the lawyer does, I want my opinion to be independent of that or of any other forensic information. Sometimes there is information in the case file that includes reports from other experts, for instance. And I don't want that to influence my results.

Bijsterveld:

One very specific question about the fillers. So the assistant decides about which fillers are used, right? Is there any specific procedure for deciding what type of fillers these are?

Schreuder:

Yes, they should be similar enough to the speakers involved. It is similar to how ear witness line-ups are presented, so there is quite some literature about that. They should be similar in, of course, age and gender. Age – approximately, of course, how they sound is not their exact age. The age of the voice, so how they sound and how they speak, so their language use should be approximately similar. A speaker on the phone speaks to someone else, so, usually, conversation partners tend to be of similar language background. Not always, but they tend to be useful fillers. We also have some sort of a collection of fillers.

Bijsterveld:

These are short fragments?

Schreuder:

Yes, the fragments are around 15 to 20 seconds. And that would be optimal. Sometimes recordings are very short so that we only have few seconds. It also depends on how a person speaks and how much a person speaks in those seconds for whether very short fragments are useful. Sometimes we say it is not enough material for doing an evidence line up. Sometimes we decide we cannot do the case, because there is too little information in the material.

Bjsterveld:

And here we talk about bias. I can see that there may be differences. It would be interesting to hear from the others whether they consider that as important.

Masthoff:

There is this issue, and I am aware of cognitive bias, there is absolutely no question about this. I respect the way you [Maartje Schreuder] are doing it when you are saying you have an assistant. I think that is a very fair way of handling it, no question. But I run into practical problems if I want to pursue the

principles. I have a practical problem on my side which is that I do not have an assistant, I do not have anybody that I could handle the materials who would be able to receive them and to organize them and present them to me in a way to be excluded of any circumstantial information. So practical issues, that is one thing.

The other thing is me as, so to speak, a solo fighter here. The amount of information that is included in the materials which are being sent to me has decreased certainly. There was a time when I had the entire file of a case on my desk. But that has been reduced. So the information I get, and I have to say even though when I had all this stuff on my desk, of course, you browse through this stuff: “Where is the CD? Okay, this is the CD. How was this stuff recorded?” And “the important intercepts are under the following number...” and so on. So here is the essential information and I don’t care about the rest what is there, whose wife was involved, and whose friend. I don’t care about the circumstantial information. It would be too time consuming, and I simply don’t care about that, family business and all. So, I do have that information on my desk, I know that there is the cognitive bias. My way of avoiding the cognitive bias is that, first of all, I simply don’t care about this. I’m too lazy to go through, to browse through the additional information. That is the one thing. More important though is that I am rather careful about applying the probability scale.

We have this probability scale. In German it would be:

überwiegender Wahrscheinlichkeit

hoher Wahrscheinlichkeit

sehr hoher Wahrscheinlichkeit

an Sicherheit grenzender Wahrscheinlichkeit

eine oder keine Stimmenidentität

So I’m really careful in entering the different levels of that scale. That may not be the perfect way of handling it, but, again from a practical point of view, at least as I perceive it, that is the only way of minimizing the risk here. That would be at this point my answer to your question.

Bijsterveld:

Do you use that scale for different sub questions or for one overall main question?

Masthoff:

Sometimes the question is: I have 20 telephone calls, 20 intercepts. Is it always the same guy who is talking on the phone, who is calling? That is the one thing and then the other would be: this guy in the 20 phone calls, is he the same guy who is talking in three other intercepts we have? And then what I do is I compare the known to every one of the unknowns. So I do not get an overall answer for 20 intercepts, but I get not one answer but 20 different answers. And each time I use the probability scale.

Braun:

I have had lengthy discussions with Maartje about the cognitive bias. I have great respect for the way they do it in Maastricht. However, I don't do it that way either for a number of reasons. The first and probably most trivial one being that nobody would want to pay for that in Germany. No German court would pay for that.

But more importantly, if we get a case, it is fair to assume it is about a felony. That is the presupposition; you do not get bomb threats or simple calls where somebody threatens their neighbour, with, for example, scratching their car, we don't get petty crime. We only get the severe cases, so you can readily assume that it is drugs or murder, or something like that. Felonies anyway. So, it would be kind of self-deceptive to say, "well I don't know what this case is about". You do know, automatically, that it is a felony case. And then very often you can pretty much guess from the materials what has happened and what it is about. I do not quite see what is gained by not knowing anything about the case.

On the other hand, I have had cases where it was rather advantageous to know certain details and to know them from the beginning. For instance, about somebody having a cold at a certain time and then it is important to know whether the recordings were contemporaneous. The questioned and the reference materials, that is. Because if they were not, then you would expect there to be a difference. If they were contemporaneous you would have to have a suspect with a cold, if you had an unknown recording, anonymous recording with a cold. This is about the best example I can think of where it is advantageous to have some further knowledge. I do not refrain from using that and even from asking for certain details. Like, for instance, if I hear somebody lighting a cigarette on an anonymous recording, I may as well ask if the suspect smokes. I would, in such a case, ask this question.

Bijsterveld:

Maartje, would you ask any of such questions?

Schreuder:

Yes, in the non-blind stage, I would want to know this kind of information. Not at the beginning. This is also incorporated in the procedure. The assistants know all this so they can, for instance, also select fragments that do not include the topics of the conversation that relate to the case. By the way, most of the things I hear in the case, but maybe because of different kind of cases, is very innocuous kinds of conversations, like "what time are you there"? "I'm now at this traffic light, I will be there", etc. And the assistant can specifically choose for that kind of fragments from the bigger phone calls, if possible. But the assistant can also decide beforehand: "Okay, I will let you do the blind analysis again, but then with this information, would that change your grouping?" And I make different groupings, also give the relations between the groupings. So, for instance, if there are two groupings that sound very similar, the only difference is that one sounds a little bit more nasal, for instance, because of a cold. I could write all these groupings, I put them in two different groupings, but they sound very much alike so they could potentially be the same speaker, but there are some

differences. So I do different forms, and then at the later stage, I could decide, oh okay this really sounds like the person's having a cold.

Braun:

But, Maartje, isn't that the assistant is then a very important person who should really be almost more knowledgeable, sorry to say that, than you are? Because he or she decides on what you get to work with and sets the stage for your case. Is that not a kind of bias?

Schreuder:

It could be, but I don't think they should be more knowledgeable than I am. But they should definitely be knowledgeable. It is also the way the NFI [Netherlands Forensic Institute] does it, that they do it for each other. They have two forensic speech experts, who do the cases for each other.

In our case, it was Ton [Broeders] and me at the beginning doing that for each other and now I have an assistant who is also very knowledgeable and trustworthy. She has done case work herself. And she can ask me questions but not specific questions. It is rather difficult, but we do manage though. Beforehand I tell her, "Select based on these and these criteria and if you have questions try to formulate them in a very general way."

Braun:

But for the assistant to ask for details such as "Does the suspect smoke" or "Does somebody have a cold?" supposes that the analysis has been done.

Schreuder:

The blind stage is not the full analysis yet, the full analysis happens when I have all information. At that stage I cannot ask for that information.

Braun:

You said that the assistant has all this information, but the assistant would have to ask for this information as well. And for that, an expertise and, in a way, an analysis is needed, because you don't know to ask whether the suspect smokes unless you have heard in the anonymous recording that somebody is lighting the cigarette. So the assistant's work presupposes analysis and yet it is meant to prepare analysis. That is the problem I have with that.

Schreuder:

There is always room for discussion, but not at that level, indeed. But I cannot always ask the police for information like that and I have to get an analysis done.

Bijsterveld:

It is very interesting, and we will see whether we have time to return to some of these issues, because I think it's important to go to the next question.

Kvicalova:

It seems that with this last topic we have been already discussing the second question, that is how your work looks like in practice. Perhaps we could then move to the third one with respect to time.

Bijsterveld:

The development of your field is very important for us as historians. Could you say something about which events or technologies were very important in your perspective?

Braun:

It is not a homogenous field, they have been and there still are varied approaches, so I am not very sure what is meant by *the* development in the field. There was of course the voiceprint issue, that you are probably familiar with, and the technologies for that. The sound spectrograph was actually a war time development. They started trying to identify German and Japanese radio operators on the submarines by comparing spectrograms and that was taken up by a chap called [Lawrence] Kersta, after the information had been declassified. But in Europe, that phase did not last long because, for various reasons, the spectrogram, or so-called voiceprints, did not succeed.

In the late 1960, the combination of audio and auditory analysis and spectrographic and instrumental analysis developed. A single person who made an important contribution was actually a psychologist and phonetician by the name of Rudolf Fährmann. He outlined what he called criminalistic phonetics. That seems to have influence at least, and that has been documented, the experts on the other side of the Iron Curtain. He is not so much quoted in West German literature, but still, much of the field emerged along the lines of what Fährmann outlined. From my perspective, and that is the perspective of the BKA [Bundeskriminalamt], it was in place by the mid 1980, because that was when the speaker ID unit was moved from the development branch of the BKA into the Forensic lab. It was made part of the forensic lab in the mid 1980s.

I think the next major development was that at that time, they tried to identify speakers automatically. There were attempts in the late 1970s and early 1980s, but they were not successful. So that was out for a decade or so and then around 2000 or late 1990s, attempts started at carrying out voice comparisons automatically. Technology changed, the statistic procedures changed, and the computing power increased so it could be done more quickly.

At present, there is a continuum between people, who – correct me if I am wrong – a continuum between people who do speaker identification work in fully automatic way around the world and those who do not use automatic comparison algorithms at all. Like for instance you are [Maartje] and I am.

Bijsterveld:

I have one specific question. What was the reason why it left?

Braun:

[It was] because the method to carry out speaker ID by auditory analysis and in a combination with acoustic analysis had been established as a sort of routine procedure, adopted as that within the BKA.

Bjsterveld:

What was the heart of the routine, what was most important in making the routine to take shape?

Braun:

That was combination of a set of parameters. In an auditory way first, and then documenting finding, doing measurements.

Bijsterveld:

Who decided that it was ready?

Braun:

The BKA hierarchy.

Bijsterveld:

On the basis of what you submitted?

Braun:

I was not there. My post was the result of that change.

Bijsterveld:

So you don't know how that happened?

Braun:

Not exactly.

Bijsterveld:

Would others like to comment?

Masthoff:

At this point we are addressing the historical dimension. From my personal perspective, I think what embraces the historical dimension is the search for an appropriate methodology in voice analysis and comparison. That is the ongoing topic or at least has been the ongoing topic. When I started practicing – that was in the late 1980s and I did this under the supervision of Jens-Peter Koester, who had already been practicing, essentially applying this method of combining auditory acoustic approach to voice analysis and voice comparison – I more or less grew into this field along with this method. And when I am saying late 1980s, that is when I started working in the field on the basis of this method, auditory and acoustic. That is when this issue of “what is the appropriate methodology?” came to my attention and has been an issue for me ever since. The automatic approach to voice comparison was not really new at that time as you already pointed out. There was the BKA in the late 1970s, but that wasn't the start either. The first attempts go back to the 1960s, late 1960s, early 1970s maybe.

And here comes another historical player in the field, from my personal perspective, and that is the Institute of Advanced Study of the Communication Process, Harry Hollien at the University of Florida. He is a phonetician, but he started experimenting with automatic or semiautomatic methods with his SAUSI system [Semi-automatic System for Speaker Identification] and long-term average spectrum. So the idea was here, it was on the table, we had the auditory and acoustic approach, but automatic methodology was already on the table and was being discussed. And that went on until the late 1990s.

That is, by the way, shortly after I became a member of the IAFP [*International Association of Forensic Phonetics*], and that later would change into IAFPA

[*International Association of Forensic Phonetics and Acoustics*], and my first meeting was in York. And there were all these phoneticians, and the discussion was a reflection of the status of methodology at that time, in mid 1990s. It is my impression that until the mid 1990s, there still was a common understanding that the method of choice was the auditory and acoustic phonetic approach. Until the late 1990s, or maybe even later, I would have to check that. And that is when automatic methodology became, in my view, a more serious issue with the emergence of AGNITIO and the BATVOX system. And that went on until the early 2010s... It became clear to me during the annual meetings of the IAFPA. That was no longer a discussion but a competition. Look, there is movement in the discussion of appropriate methodology and along with AGNITIO and BATVOX, the auditory and acoustic phoneticians had to face this discussion concerning the appropriate methodology in a serious manner more so than before. So the phonetic approach and the machine automatic approach were then competing and it looked like the machine approach was gaining a momentum, starting in about 2005 on.

So this is an ongoing debate about appropriate methodology. And this has been with us for well nearly 20 years now, off and on, automatic versus phonetic approach. And, as far as I can see right now, this idea, this machine approach or automatic approach, is on a ballistic curve, so to say. First preparations in the late 1990s and then once we get into the 2000s following the millennium, the ballistic curve starts and has seen an apex in 2015 or at about that time. And it seems to be coming down a little bit in a sense that all colleagues I know, who have been in both areas, really, phonetic as well as machine, are beginning to realize that the machine approach may after all not be the answer to what's the best practice in voice comparison.

The main argument was that the phonetic approach has a strong subjective element, and we need to take care of the cognitive bias, of course, as I pointed out. So how do we get rid of cognitive bias and how do we get rid of subjectivity in the method? And this is, of course, a good point. The machine or automatic approach seemed to be a solution, but it seems that the proponents of the method are beginning to realize – to come back to the point – that the machine approach is not actually the solution. That the machine approach or the automatic methodology may support or may be a supplement to the phonetic approach and not more. And that the machine approach, at any rate, requires to be supervised by the human expert and by the human auditory perception.

We are in 2020, we have the ongoing search for the appropriate methodology. The auditory phonetic method is still around, and according to [Peter] French and Erica Gold⁵... I don't know, it has been a couple of years now, but they had this questionnaire going around in...?

Braun:
2019.

⁵ French, P. & Gold, E. (2019). *International practices in forensic speaker comparison: Second survey*. *International Journal of Speech, Language and the Law*, 26(1), 1-20.

Masthoff:

It seems that for most experts, in spite of everything that happened, the gold standard, so to say, is still the auditory phonetic method.

Bijsterveld:

Even among the more acoustically oriented ones? You said that even in the machine expert scene there is an increasing doubt now.

Masthoff:

I may be wrong, but that's my impression.

Bijsterveld:

What kind of things do they refer to then? What are the specific issues of which they are critical of themselves? What can the machine approach *not* do?

Masthoff:

The problem of the machine approach is, and Angelika has a much better scope of that, is to deal with adverse conditions in recordings. Different digital formats, different length of recordings, background noise, background population. Angelika, I think you would be the better person to answer that question.

Braun:

I am not as convinced as you are that the people who are advocates of the automatic approach or machine approach have that many repercussions.

Masthoff:

I'm sorry to interrupt, that is my impression as far as, I would say, Central Europe is concerned. I don't have that scope what the situation is on an international basis.

Braun:

There is also the Interpol survey done by [Geoffrey Stewart] Morrison⁶ where the wide range of methodologies become evident.

I'm not so sure that those who advocate the machine approach are that critical of themselves. It's primarily phoneticians, not so much computer people, who have doubts. And it is forensic practitioners, who have doubts. That is, I think, almost the more important line of divide between those who are sceptical of automatic approaches, and those who are not. There are very few people who regularly do forensic work who are very and truly convinced of automatic approaches.

There is a lot of research being done, so you get those equal error rates, and you get other ways of measuring the quality of the different automatic systems. That of course sounds very good, but then the forensic practitioner comes in and says: "Hey what about if my suspect whispers?" "Oh, no, that we can't deal with." "What about if I have only three seconds of material?" "Oh no, that will not work." "What about if the offender screams and shouts in the middle of the

⁶ Morrison, G. S., Sahito, F.H., Jardine, G., Djokic, D., Clavet, S., Berghs, S. & Goemans Dorny, C. (2016). INTERPOL survey of the use of speaker identification by law enforcement agencies. *Forensic Science International*, 263(June), 92-100.

recording?” “No, we cannot deal with that very well.” It is, I think, and I may be wrong, that the phonetically trained practitioners were more sceptical about the automatic approach than those who have tested it under laboratory conditions. Which is all fine, but it has just not performed. The latest case that I cited, where there were three previous reports, incidentally two of those previous requests were requests that also contained the machine approach and they were by no means unanimous. So, it is, from the practitioners’ point of view, an illusion to think that two machines always come to the same result. It can be held against phoneticians that they do not always come to an 100 % identical result, but the same is true for machine analysis.

Masthoff:

I think that Angelika raised a very important point here, and that is the practitioner fraction and the academic fraction. A very important distinction to make in all this. The competition between the phonetic and the machine is, indeed, not only a competition for best practice but also between the academic and the practitioner fraction. It is a reflection of the discussion between the academic and the practitioner side of the field.

Bijsterveld:

You said that practitioners are usually phonetically trained, is it always the case?

Braun:

This is the case in Europe, I would say. By and large.

Masthoff:

Yeah.

Braun:

But, of course, there are countries, where for instance voiceprint analyses are still being carried out and we would not consider the people who do that to be phonetically trained. So it depends a little bit on the part of the world that you are looking at and I am afraid that in the US there are also some people who are definitely not phonetically trained, who carry out forensic work, and they are not in the machine fraction either.

Masthoff:

Two or three weeks of special training and you are good to go.

Braun:

Yeah.

Schreuder:

There are also some different parties in Europe. I think in Austria there was this...

Braun:

Yes, you are right.

Schreuder:

...someone who studied chemistry and does all kind of forensic work

Braun:

.. she has a degree in chemistry. I totally forgot about that. A private lab.

Schreuder:

She was present at the latest IAFPA meeting.

Bijsterveld:

But I guess that the courts decide about whether they are accepted as experts. How does that work?

Braun:

It worked despite a very very vigorous and very critical report. The court said, the particular person has been accredited by the local chamber of commerce, and so she is considered an expert, period. The courts were totally oblivious to, or ignorant of, the criticism. It was no success whatsoever, even though I found the arguments very good.

Kvicalova:

How common is it that more than one expert is invited before the court? Is it usually that experts are to represent these two different approaches, that one is the phonetic expert and the other one would be doing automated analysis? Or is this not a standard?

Braun:

It can be either way. It is not that experts are chosen based on different methodologies. One of the ways of getting in another expert is if the court is not satisfied with how the results are presented by the first expert. If the court is not convinced, then the court will appoint a different expert. And in very rare cases, the defence gets another expert in, because they have some objections to the first report. Maybe you [Herbert] can say more about that. You have also experience with being called as a second expert.

Masthoff:

The challenge really starts for the judge in how he goes about selecting an expert, to begin with. What they usually do is that they *google*. It beats me how they finally come up with the expert. What happens though is that judges tend to call in experts they have heard about from colleagues, from talking to colleagues, going to meetings. Or by simply hiring the expert they had in the cases before. Or they called the BKA. That is what happens. That is one of the approaches. They either google, or they call the BKA: "I have a problem here, I need a voice analysis. I need a voice comparison; can you do this?" If the BKA says: "It cannot be done within a reasonable amount of time." They refer to others, experts they know of, so they make the recommendation, or they don't make a recommendation, but they have a list of experts they offer: "Here you go, take your pick."

Many times, it is simply by the gut feeling the judge has, "it sounds good, what I read or hear, so I go for that person."

Bijsterveld:

I guess it is slightly different in the Netherlands?

Schreuder:

Yes, we have the register, but this field is not in the register, so it works more or less the same, I think. But there is also the Police Academy, as the *landelijke deskundigheidsmakelaar* [national expertise broker]. There we are registered. But it is also that they check expertise in a way, so they have peer review.

Bijsterveld:

So it is not that they just say: “Okay, we have two forensic Institutes in the Netherlands, and we simply get the experts”?

Maartje:

More or less, yes.

About the gut feeling: I don’t know how it is in Germany, it may be different, but my feeling is that most judges and also prosecutors and others in the legal arena, don’t even know that there is such an expertise. So often it is the police officers themselves doing voice recognition or identification. They ask interpreters to do it, and, sometimes, even audio technicians.

Braun:

But I have not come across that very often. It may happen, but I don’t think that I have ever...

Masthoff:

...At least not in Germany.

Braun:

They will approach the state forensic labs or the federal forensic labs and ask for help. That is what the police would do in Germany. They would not ask an audio technician or an interpreter.

Schreuder:

That may happen more with defence lawyers, who approach audio technicians. But the police do that all the time, the policemen interpret it all the time while they are listening to the taps, they write down “we recognize this voice and that voice.” And it is usually only when the defence lawyer says “No, that is not my client” then an expert is called in.

Braun:

In Germany they would do that as well, they would try to identify speakers. But they know that for a trial, they need an expert.

Schreuder:

And for us that is, I think, more often the case. So that is a difference between the two countries.

Kvicalova:

I have a question about those two types of expertise, the automated or mechanical, and the phonetic. Is it sometimes the case that one expert combines both approaches, or is it completely rare? Because it seems like an ideal solution to combine both, sometimes it could be useful.

Braun:

It does and there are people who do it. However, what do you do if the two contradict? Who wins? That is an answer which is usually not easily provided by the people who do use both. What do they do if there is a conflict.

Schreuder:

This is also the topic of discussion and debate in the conferences we attend.

Bijsterveld:

You could imagine a situation in which you use the automatic analysis for particular issues, but not for others. That is at least what I get from the literature.

But I have one other question, perhaps a final one on the machine issue. Of course, you are experts in identifying voices and sounds, but what about the *detection* of sounds, so groups or sounds, like: “this is a car”. Even though you could say “okay that is just what I do by listening to it.” But there is a lot of detection software that is able to say: “this is a broken window”, or a particular group of sounds, not voices but sounds. What you think about that? Is this even relevant for your work?

Braun:

I do not know of anybody who has ever tried that, probably because the quality of those recordings is so bad that a machine would really not have a chance. There is of course software that tries to detect foreign accents or certain regional accents, but those machines are very good if they can tell a Chinese from a French accent in English, but any expert can do that in two seconds, you don't need a machine for that. It is about distinguishing a Serbo-Croatian from a Polish accent, that is where the hard bit begins, in distinguishing that.

Masthoff:

I think in detecting sound, or types of sounds other than voices, I don't think that a phonetician has any advantage over a layperson. Breaking glass...

Bijsterveld:

But that is a kind of software that is used for monitoring, not for identification.

Braun:

If it is at a very low level, the lay person probably wouldn't even hear it and the phonetician does hear it in the background. That is the difference.

Bijsterveld:

And why do you hear it?

Braun:

Because we listen for it. You listen differently, you don't listen for just speech, but you listen for everything that is happening on the tape.

Masthoff:

Yeah, I would agree on that.

Bijsterveld:

Do you feel that it developed over time? Looking back, do you think that you became better listeners?

Braun:

I don't know. The only way to find out would be to go back to the tapes I listened to in 1986 or 1987 and then find out whether I can now understand more or detect more. I don't know if I have become better.

Masthoff:

The phonetician doesn't have a better sense of hearing than a lay person, but the advantage of the phonetician is that he has an analytical approach to hearing or listening. An advantage over the lay person. That is particularly important when it comes to the analysis of voice and speech. So the detection in the sense of having an analytic approach to listening to the tapes is the advantage, but *per se*, telling the difference between a motorcycle going by or breaking glass is nothing that a phonetician would be able to do better than a lay person.

BREAK

Bijsterveld:

I would like to come back to what you said about Fährmann and that he was not cited so often in Germany.

Braun:

In West Germany.

Bijsterveld:

He might not have been so often cited in East Germany either.

Braun:

Yes, he was. There was a striking difference.

Bijsterveld:

Yes, by [Christian] Koristka especially, who was of course dominating. At least in the GDR. I think it may be different for Czechoslovakia.

Kvicalova:

He was cited somewhere, but he was not a famous figure there.

Bijsterveld:

When we do research, we read a lot of those publications from that time so we come across these names anyway, but can you mention other people who would be important in West Germany at that time? In the 1960s?

Braun:

I wouldn't know. Even in the publications that were there, I cannot think of any Germans cited. There were of course the studies from the US, but I'm not aware of any West German publication on the subject.

Masthoff:

If I'm not mistaken, until the late 1960s, there weren't any relevant publications specifically to the issue of speaker identification.

Braun:

Except for Fährmann. That was in 1967, that is why I remember it.

Bijsterveld:

I found a lot of references in Koristka's work, also to other [West-]Germans. They may not always have been very specific on speaker identification, but they were concerned with all kinds of other issues like using the magnetic tape recorder, the authenticity of tape, whether or not something was relevant and trustworthy enough to bring into court.

What I could do... At home, I have a long list of those. I made a specific list of where he [Koristka] refers to other types of languages and countries, which I might share with you at one point to see whether you recognize any of those. It would be too detailed now, but I would be interested to see whether you would recognize some of these people as being of any relevance later.

Masthoff:

Did you come across Wiktor Jassem? And some colleagues in Poland.

Bijsterveld:

Anna is more knowledgeable about Poland.

Kvicalova:

It actually seems to me that nobody is very knowledgeable about Poland. It is said that they established some sort of program devoted to voice analysis already in 1963 or in the first half of the 1960s, so I would like to know whether you have and references on that.

Masthoff:

I would have to check.

Kvicalova:

No one has really work on that. There was Stanislaw Blaskiewicz, who was the most prominent Polish scholar working on voice identification, but we don't know much about how the institutional background developed, when a department was established. What were you saying about Wiktor Jassem?

Masthoff:

Yes, he did some work there. And it's not a coincidence that The University of Florida, that I mentioned earlier, they had a number of joint publications. Harry Hollien along with some Polish researchers, in the development of the semiautomatic speaker identification system.

Bijsterveld:

Do you know anything about how that connection was established?

Masthoff:

Not offhand.

Bijsterveld:

Is he still alive? [Hollien]

Masthoff:

Yeah. I haven't heard anything else. He is 96 or so.

Braun:

He was still alive half a year ago, I had some exchange not with him personally, but with Ruthie [Ruth Huntley Bahr].

Bijsterveld:

But what do you know about the relationship between Western... or what kind of experience did you have with the non-Western experts before the Wall fell, or even after that? Can you say something about it?

Masthoff:

As you yourself [Karin] have published on the *Staatssicherheit* and speaker identification, I am in contact with two people, who worked for the *Volszpolizei* and did speaker identification and voice analysis there.

Let me put it that way, I have no mandate to elaborate on that at this time for the simple reason that we have a little piece of work in progress, in preparation, to shed some light on the *Volkspolizei* speaker identification. That is between me and a number of other authors. The thing about this is that it is a little bit a matter of trust here that is involved to exhibit all of the information. Whether there is good reason for that or not that is a different ball game. Basically, these people are ready to talk about it. I can say that it is no "big secret" involved or anything of that sort, but simply a matter of being prepared to talk about it.

Bijsterveld:

That is also news, that there will be a publication.

Masthoff:

That the story will be told.

Bijsterveld:

Will you have interviews with them?

Masthoff:

Interviews indeed. I have started some of that.

Bijsterveld:

Can I ask why you started it? Because that is intriguing.

Masthoff:

The reason for this is that in what is being or what has been published, there is still a lack of information on the work that was done in East Germany. I don't know about Czechoslovakia and Poland. There must have been, I suppose, some relation or cooperation between the Poles and the GDR. And the work they did, that is my impression, was highly qualified at the time and throughout. You could discuss, and this goes back to Fährmann a little bit, about making psychological assessments and all that. But apart from that, from the strictly phonetic point of view, they did a pretty good job. The point is to be able to tell about the full scope of the development of speaker identification in Germany by at what more or less developed simultaneously in the East and West.

Even though the one side didn't know exactly what the other side was doing, they were similar developments from a methodological point of view. We had talks before and we had meetings and the idea was to tell the full story. What created the impulse and let us put the information together.

Bijsterveld:

What time frame?

Masthoff:

The time frame would be late 1960s, mid 1970s and on. Especially the 1980s, though, late 1970s into the 1980s.

Bijsterveld:

We had different parties here [East Germany]. The Humboldt University and its Criminalistics department with some experts, mainly Koristka, working together with others; students at the *Juristische Hochschule*, the *Stasi*, different departments within the *Stasi* with different archives; and the *Volkspolizei*. So you are now flying in from the *Volkspolizei*?

Masthoff:

Right.

You were asking earlier what happened at the time of 1989. There were meetings in the BKA, I am not able to remember those names. There were two guys... [anonymized]. Two older guys, who were specifically trained or qualified in dialects.

Bijsterveld:

There were definitely people specialized in dialects, because in the *Stasi*, I find endless documents on dialects.

Masthoff:

That was a big deal in the *Stasi*.

Bijsterveld:

It became more and more important.

Masthoff:

Yes, being able to locate people on the basis of their dialectal features.

The two guys were there only for one meeting and after that they were gone.

Braun:

They were probably the victims of [Joachim] Gauck.

In 1989 I was still working with the state lab in Northern Westphalia. In that capacity we of course could not have any contact with our East German colleagues at all. I wasn't even sure whether such a unit existed, so I knew absolutely nothing. There was no exchange whatsoever. And it couldn't have been either, because it would have been considered spying, or at least an illegal contact with the enemy.

My memory only goes back to the time after 1989. That is when I visited the Volkspolizei lab in East Berlin. And three of their experts also came to see me together in Düsseldorf with one former *Stasi* expert by the name MH [anonymized]. It was not his real name, of course. And it was striking that the Volkspolizei experts were treating the *Stasi* person with great respect and the *Stasi* person was quite disrespectful of the Volkspolizei people. There was apparently a clear hierarchy between those two institutions. I could tell that quite well.

Then we had further contact with the former Volkspolizei experts, one of whom also, I think, joined the IAFP and came to their meetings.

After the reunification, when we were talking about the methods, it turned out, as Herbert Masthoff mentioned, that the developments had taken a similar path. Except that the colleagues from East Germany, following Fährmann, placed the greater weight on psychological analysis, which is what we did not do at all.

Bijsterveld:

That being that the voice expresses particular character traits?

Masthoff:

Not only character traits, but also evaluating the emotional states.

Bijsterveld:

Actually, the emotional states seem to be very important, they returned in all documents of the *Stasi*.

Braun:

Yes, but that was *Volkspolizei*, not the *Stasi*.

Bijsterveld:

But of course there were contacts between them.

Braun:

Yes, but surprisingly few, from what I gathered. There was not a mutual exchange, because certainly the *Stasi* did not consider the *Volkspolizei* of being on even terms.

Bijsterveld:

Koristka was also very critical.

Braun:

I don't know, I never met Koristka.

Bijsterveld:

In the sources he is complaining about the police as being not skilled enough. That is what he said.

Masthoff:

And vice versa.

Bijsterveld:

That is interesting and it is great that you are going to publish about it. Just to make it explicit: we as historians do not say who is right, but what others say is true, and we also try to track where they get their information from.

Braun:

Well, and in conversations it emerged that at least the *Volkspolizei* experts had the closest ties with Poland. Yes, that is what they regularly mention. More so than with Russia.

Bijsterveld:

In the Stasi documents, the Stasi itself actually complained about the KGB not giving as much information as they had hoped, and also having the feeling that they did not get the newest instruments, the newest versions.

Kvicalova:

It was the same in Czechoslovakia. It seemed that they were rather disappointed, they didn't really know what was happening in the USSR in terms of voice analysis. They [the USSR] did not really share that information.

You [Angelika Braun] said that it is often mentioned that the Germans cooperated with Poland. Are there any references which could tell us more about with whom they worked, if we wanted to look more into it?

Braun:

I wouldn't know.

Bijsterveld:

The sources, but also the reviewers of our work (very often we do not know who these people are) refer to Poland. The sources say that they did visit Poland or that the experts had conferences...

Kvicalova:

There is a reference about a conference in Kiev in 1980, in Ukraine, but we have no information on that only that it dealt with speaker identification. It is all very obscure and sometimes we have only bits and pieces of information.

Bijsterveld:

We would really like to know more about these relationships.

Braun:

I remember there was a delegation from Poland visiting the BKA in the second half of the 1990s and that they were working in speaker identification. Four or five people, but I do not remember the names and much of the talks. We as experts didn't get to talk to them for very long and I am afraid I don't remember anything. I wish I had taken notes.

Bijsterveld:

There might be some traces in the BKA archive then. We just need to trace back those names. That is already very helpful.

Masthoff:

Besides Wiktor Jassem, it would be worthwhile revisiting the papers by Harry Hollien and Wojciech Majewski.

Braun:

And there was a third one...

Bijsterveld:

We will find that, that is no problem. I mean I have huge files...

Masthoff:

That may exhibit the beginnings of the cooperation between the University of Florida and whatever that institution was in Poland. Jassem is a phonetician, by the way, but Majewski is an engineer. That is the interesting part here. And the partner of Hollien, who is phonetician working on the development of an automatic or semi-automatic system, calls in the expertise by Majewski from Poland. So that makes sense, in a way.

Bijsterveld:

It makes sense but it is also highly intriguing. How did they work and where were these people in the US, for instance?

Masthoff:

What I do recall is that Majewski, for example, spent a term or so at the University of Florida.

Braun:

Is Janusz Zalewski the third name?

Masthoff?

Beats me, it could be.

So there was this interchange of “Wissenschaft Austausch”, something along that line.

Bijsterveld:

But you wouldn't expect that. So that is interesting.

We do know, of course, that the KGB knew everything that happened in the *Bell labs*. It is more or less clear to me, given the fact that they sort of copy-pasted what happened there, although it was highly classified. That is intriguing as well.

Braun:

They also built a spectrograph and it was happening after it had been developed by *Bell labs*. They must have had gotten hold of the documents.

Kvicalova:

The knowledge also spread because of Kersta's publication.

Bijsterveld:

But that was much later. They already copied their experiments in the 1940s. And it looked so much the same that I think.... It is likely they had their spies there. The only thing you do not know is who it was.

The topic of Poland kept returning in reviews, but even reviewers are vague about it.

Kvicalova:

I have the same experience. The reviewers always ask me: “why don't you explain more what was happening in Poland?” But they don't give you any reference.

Bijsterveld:

Someone gave a reference to a French source, I went through it back and forth, but there was no helpful information.

Braun:

Are there no Polish experts cited by Koristka? I remember seeing names and references in those publications.

Bijsterveld:

Yes, definitely. The problem is that I can't read those publications. That is actually how our collaboration [with Anna] begun, because she can read Czech, but also a bit of Polish. We met at the Max Planck Institute for the History of Science, and I asked her "Do you know anything about that?"

Kvicalova:

And I did not. It just seems that there is much more to be discovered and that the language barrier is also part of the problem. The thing to do now is probably to go to Poland and get to know more Polish scholars and former Polish experts. I am sure that we will find out more about what was happening in Poland. It is interesting that their research really started already in early 1960s and that it was high-quality research, from what I gathered. There are so many similarities between what was happening in the West and in the East, but we do not really know much about the nature of their contacts, whether they were any or whether it was mostly an espionage.

Perhaps another interesting point of contact would be the transfer of technologies, because they definitely did cross the ideological borders, at least from the West to the East. In Czechoslovakia, they had a so-called *voiceprint* machine in early 1960s, which came from the US, and it was not just that. Technologies crossed the border, but I think that it was mostly from the West to the East. In Eastern Europe and in Russia they had all those dreams of developing new splendid espionage technologies, but they were not that successful actually.

Braun:

I know that the spectrograph made its way to East Germany despite being on the CoCom [Coordinating Committee for Multilateral Export Controls] list, under guise of medical equipment. That is how they got it in there.

Masthoff:

Two more names have just occurred to me in the context Poland. One of them is Czesław Basztura and the other guy is Ryszard Gubrynowicz. Engineers, but with the inclination to phonetic applications. But how they relate to Majewski or Jassem, that I don't recall.

Kvicalova:

It is interesting that you said that they were engineers with an inclination to phonetics. That is also my experience with the Czechoslovak experts in the 1970s and 1980s, that they combined the expertise. They were engineers, but then they got some basic education in phonetics as well and that created the expertise from the scratch. This is why I asked the question in the beginning of the workshop whether it is still the case nowadays, or whether such combinations have become really rare.

Bijsterveld:

Koristka had also a very mixed pack of knowledges, so to say.

It is good to see how you [Angelika] look at it, because you know his main publications.

Braun:

I think that the general impression of Koristka was that he was not an expert in anything, but he could put information together, and publish it. As we indicated earlier on, this sort of disrespect was mutual.

Bijsterveld:

The former *Stasi* expert you talked about was Koristka?

Braun:

No, that was somebody by the name of MH [anonymized] He was the only *Stasi* expert who made it into 1990. He was still being employed and got to Düsseldorf to visit the state lab there, but then, after a couple of weeks, he was gone. He did not pass the Gauck examination.

Bijsterveld:

I did speak to Koristka's *Promoter Vater*, that is a signal processing expert, but we have never been able to retrace Koristka's PhD thesis. It was in the archives, it should have been in the library of the Humboldt, but it has disappeared. Many people are looking for it, but nobody can find it. The question is "Why has that disappeared?"

Masthoff:

It is apparently not the only thing that has disappeared or is being kept elsewhere.

Bijsterveld:

I am in contact with another person from the Criminalistics department because he actually took some of the archive of the Humboldt Criminalistics department simply to his home. So much happened of course in these early days after the *Wende*. Sometimes people made things disappear, but some also took them with them. And Koristka does not want to talk anymore, I have tried. It is my own mistake, I have tried twice and the first time he was willing, and then when I actually wanted to do the real interview, he had changed his mind. He is still alive as far as I know.

Perhaps you [Angelika] have more success. Or not, he does not want to talk about it?

Braun:

No, not to me. I think it is probably for similar reasons that the former experts won't talk to me anymore, because I was part of the official lab at the time. There is not enough trust there, no.

Kvicalova:

Maybe you should send him your latest paper, that might convince him?

Bijsterveld:

I do not have his address, because he kept that secret. I only have a phone number.

WE [anonymized] has talked to him. He is not a historian, but someone who is doing media studies, who did establish the contact with the *Promoter Vater*, and also with Koristka himself. So I had the phone number from WE [anonymized] [...]. He [Koristka] just said "This is the past, it is all over, *lass mich in Ruhe*."

Kvicalova:

If I get it right from what you have described, there was a great discontinuity in terms of personnel in the 1990s as most people who worked in speech identification before that did not pass fast the Gauck or just stopped working there. Or was there still a considerable number of experts who just continued doing the job?

Braun:

No, all the *Stasi* experts failed to pass the Gauck procedure. I think that there were three experts from *Volkspolizei* who went on.

Bijsterveld:

I gave one lecture in Berlin in 2009, I think. It was my first lecture about this, and I had a strong feeling that there were experts in the room. I could see them and they may have been the *Volkspolizei* people. Someone from the *Stasi* archive did contact me afterwards and she helped me a lot, but these people remained in the background. I was very close to just going forward to them and asking, but it also seems a bit.... You don't do that so easily. [...]

Do you have any idea when your paper is going to be published?

Masthoff:

Towards the end of the year.

Bijsterveld:

Is there anything you would like share with us, either physically, such as articles, or anything that you think we should know about your field that you find important.

Schreuder:

There are some publications by *Interpol* that give the state of the art of the field. Especially the one that Ton [Broeders] wrote in 2011.

Braun:

Which one? The *Interpol* survey was in 2017.

Ah, that was the report he did for the *Interpol* meeting in Lyon.

Schreuder:

Yes, and it gives a very good overview. But when was that?

Bijsterveld:

Ton [Broeders] has of course also published about the history of the field, and that was one of the first publications I had seen at that time.

I have a question concerning the associations. We have the forensic phonetics and acoustics, and also the association of forensic linguists. Can you say something about the most important differences there? Or are these completely different fields?

Braun:

The difference is quite simple. The forensic linguists are concerned with texts only and the phoneticians, of course, with the spoken word. They developed more or less in parallel, but they still share the journal. But they are now completely separate.

Schreuder:

Yesterday, I browsed quickly through my bookshelves. I inherited many journals and books et cetera from Ton and I found something, but it is not about forensic phonetics only, but the history of American phonetics. Some chapters might be interesting.

Braun:

If you digitalize it, can I have a copy?

[Karin is browsing through the book and reads aloud some of the chapter's names]

Masthoff:

Another important name with an impact on the discipline as far as voice analysis is concerned is, in my view, John Laver.⁷

Braun:

What about Francis Nolan.

Schroeder:

Phil Rose.

Masthoff:

...at a later stage.

Bijsterveld:

But of course, this is all later than the research we do at the moment. But it is important for us to see how the field developed and to put our research in perspective.

Braun:

Francis Nolan's dissertation is from the 1980s.

Bijsterveld:

Voice identification was a big thing in the UK.

Masthoff:

I suppose it is no coincidence, that the IAFP started in York in the UK.

Braun:

Peter French started the initiative.

Masthoff:

The British had a strong impact on the further development of the phonetic approach to voice analysis and voice comparison at that time, that is the 1980s and 1990s.

Kvicalova:

There is one question I would like to go back to and that is the notion of objectivity. We partly discussed this in the beginning, that is, how you achieve objectivity by, for example, not knowing anything about the case in advance or how you deal with the subjective assessment and subjective listening experience. How did the way you talked about objectivity at the time you started in the 1980s and 1990s changed over time? Was this even a topic you came across in the beginning of your careers? Was it being discussed?

⁷ Laver, J. (1994). *Principles of Phonetics*. Cambridge: Cambridge University Press.

Where my question comes from – in 1980s Czechoslovakia, for example, the notion of objectivity was often mentioned as a selling strategy of this new field of expertise, and it was usually connected to the high hopes put into the prospect of automated speech analysis. It was also connected to the historical perception of sonograms, which were seen as a new and more objective method. On the other hand, they still relied heavily on listening and auditory perception. Were there any shifts in the field with respect to how objectivity was understood?

Braun:

I think we should distinguish between objectivity and impartiality. To me, impartiality represents the inner motivation of the expert, and that issue became very obvious to me when I worked on the Rodney King case. In the United States, the system is such that as an expert, you have a very difficult time to maintain impartiality, because that is not what is required. It is required that you are an expert working for one side and that was something that I was definitely not used to in Germany. Even at the BKA, we, as BKA experts, sometimes got criticized for being sort of police experts. But, at the BKA, the Forensic Institute was totally remote from the police part of the BKA, it was another unit. We were treated differently and nobody, thank goodness, nobody at the BKA has ever tried to influence us. That is impartiality, which I was granted in Germany, but which you cannot maintain in the US system. Or not easily maintain because of the constraints. For instance, you are not allowed to state all your findings. You can state only the findings which your side wants you to state and then it is “okay thank you”. So that is one thing.

Then objectivity. That is the thing about “human” versus “machine”. “Human” always has the connotation of being subjective, “machine” has the connotation of being objective and I think that is very wrong. Of course, there is a subjective element in anything that humans do. And you are probably trying as hard as one can to work around that, but what I emphatically dispute is that the results of the machine are objective. I would simply dispute that because there is always a person operating the machine. That person makes a number of very important decisions, like for instance which materials to use, which reference population to choose and so on and so forth. The so-called objectivity is a bogus objectivity, and the machine is no more objective than the person operating it.

Masthoff:

It starts with the design of the machine. At that point. You have a bias at that point already.

Bijsterveld:

But in our field, this is more of an ideal, it is referred to as the ideal of objectivity. It is not that we say: “it is objective”. But it is often used rhetorically in many ways.

Braun:

It is often said by the US experts. “This is a disadvantage of your method; you are subjective, and the machine is objective.”

Schroeder:

Another important part of objectivity is whether the measurements are objective. For that it is also very important to have big databases and that is still really a problem in our field. Peter French also mentioned that a few years ago during an IAFPA conference. That this was already mentioned at the start of the IAFPA that we should build bigger databases to have better population statistics. And that has still not happened.

Bijsterveld:

That is, having voice data or phonetic data across all kinds of groups of people?

Schreuder:

But also representative. It means not only high-quality standard dialect or standard accent of people with higher education. This is very difficult, the problem of what choices you make, or you do not make to build such a database. There are some databases, but they are rather small. The NFI has now a database of forensic material, I think that is an improvement. It is only for internal use, I think.

Braun:

It is very difficult to have everything you need in your database, because you may need a database of, I don't know, people with Hebrew as their first language, Arabic as their second, and speaking Dutch.

Schreuder:

When he was 18, and when he is 28, and when he has a cold, when he is drunk, in the morning and the evening.

Bijsterveld:

You have a super interesting field.

03

TRANSCRIPT
OF THE INTERVIEW
WITH PETER FRENCH

TRANSCRIPT OF THE INTERVIEW WITH PETER FRENCH

Online, via Zoom, 9 November 2022

PARTICIPANTS

Chairs:

Prof. Dr. Karin Bijsterveld

Dr. Anna Kvicalova

Participant:

Prof. Dr. Peter French

Bijsterveld:

What was the most memorable speaker identification case you worked on and how exactly did you use your expertise in that particular case?

French:

Does it have to be a speaker identification case? Because I noticed for example in Angelika [Braun]⁸'s responses, she talked about the Rodney King case. I think that from my point of view, some of my most memorable cases were not speaker comparison cases, they have been other types of forensic phonetics, and forensic audio analysis.

Bijsterveld:

That's also okay because for us that's interesting as well.

French:

One of the problems about addressing this is that cases are interesting from different viewpoints. Some of the ones that are most interesting from a public angle, are not terribly interesting from a technical phonetic or acoustic angle, and vice versa. Some of the most interesting cases I've been involved in from a technical viewpoint have been rather trivial crimes, which perhaps only make it into local newspapers, whereas the very high profile ones – for example, I was involved in the case against Slobodan Milošević in the International War Crimes Tribunal⁹ in the Hague, and I've been involved in another very high profile things such as the [George] Zimmerman case¹⁰ which I think was referred to by Angelika [Braun] – they were not terribly interesting technically, but of course, they are the ones that people remember because they are all over the newspapers and the news media of the time.

I guess when you look at it from a public angle then yes, I would say the prosecutions of Milošević and the various generals in the Balkan states in the war crimes tribunal. I acted for the United Nations against general Obrenović. I also produced evidence in the case of Radovan Karadžić, but that evidence didn't go to court. And, of course, Milošević himself. In that particular case of Milošević, I wasn't employed by either the prosecution or the defence, I was appointed by the court. So as a witness who was independent of prosecution and defence.

I guess, those cases were also very memorable, because they were very harrowing. They generally involved interceptions of radio communications between the various generals, and these were communications which involved instructions to eradicate villages and all the people in them. You know, 'leave no one alive' and things like that. It was very trying, I guess, emotionally. No matter how one tries to distance oneself from that aspect of the material, nevertheless, that does creep in.

⁸ One of the participants in the witness seminar on speaker identification, 17 June 2022, see above.

⁹ International Criminal Tribunal for the former Yugoslavia.

¹⁰ *State of Florida v. Zimmerman* (2012), see footnote 4 above.

Other cases that I've been involved in on that sort of level have been international commissions of inquiry. I was involved marginally in the inquiry of the Boipatong massacre in South Africa.

As I said, none of them was particularly memorable from a technical viewpoint, it didn't bring particular challenges. In the case of Milošević, my work simply involved collaborating with another member of my firm, a sound engineer, and also with speakers of BCS, Bosnian Croatian, Serbian, and examining the recordings for any evidence of them having been edited or tampered with, and it was a pretty straightforward case.

Bijsterveld:

So it was really about the authenticity of the recordings themselves?

French:

In that particular case, in the Milošević case, it was about the authenticity of the recordings. What we were looking at there was whether these were continuous unaltered records of what had been said, or whether material might have been removed from them or material interspersed or interjected, and that sort of thing. In the cases of the other generals, the work was speaker comparison, or as some people call it, 'voice identification'. And again, these were pretty straightforward, except in Serbian as it was a language that I was not familiar with. In order to do the work in a case like that I need to have available to work alongside me somebody, who has a background in linguistics and preferably phonetics, as well as being a native speaker of the language. So they were a little bit interesting from that viewpoint, but not terribly interesting.

Bijsterveld:

Is that expertise then combined in one person, or do you have someone who is native in a language and someone who is a phonetician?

French:

I am the phonetician, and I would require them also to be a phonetician or at least a linguist with some training in phonetics as well as being a native speaker.

I think it's particularly important in the speaker comparison cases because one of the things one's looking for when comparing the recordings are features of a personally distinguishing nature, so things which are not just accent or dialect features, but things which mark out the speaker from the general speech community from which they have come. In other words, the idiosyncrasies. And in order to determine what is an idiosyncrasy one needs to know what the norm is, not just for the language in question but preferably for the particular dialect as well. And that's why it's important to have somebody who is both a phonetician and also a native speaker working with me. I mean, the problem is that you could form a conclusion which was based at least partly on what you saw as a strange pronunciation of let's say the consonant "L" and then it might subsequently turn out that it's not in the least strange for the particular dialect of Serbian, that it's pretty normal. So it really is important to have somebody like that to work with me.

Memorable cases from a technical viewpoint, which were not so serious: one concerned a case which is very well known certainly within the UK, but maybe not internationally, which involved cheating on a TV game show. It was called “Who wants to be a millionaire”. What happened in that particular case was that there was a participant, an army major at that time in the UK, who was a competitor, and he was answering multiple choice questions. And as he was answering the questions, he would mull over, talking aloud what the potential answers were. A particular question might be something like: “What is a homburg?” And the four choices might be: a type of shoe, a type of overcoat, a type of jacket, or a type of hat. And he would mull over that question by saying: “is it a type of shoe?” “Or is it, I don’t know, maybe it’s a type of jacket?” “Or it could be a hat” And of course, it’s a hat. And when he would produce the right answer, there would be a cough came from the audience, which was obviously someone prompting him.

I was approached by a police force in the UK because this guy won a million pounds, and the TV company was very suspicious about this coughing and whether or not he genuinely won it, or whether there’d been some cheating going on. The question was: “Can you examine the recording and listen to the coughing and tell us where in the audience this coughing might be coming from?” The problem was that there was an auditorium which was circular, different levels of seating, so on and so forth, and I think – I’m shooting a little from the hip here – but I think there was somewhere around about 230 or 250 people in the audience. And of course, when we got this recording, my reaction was: well, I don’t think there is really a hope to find out where the coughing is coming from. But then, as we got into the case in more detail, it became clear that the recording we had been given was a mix. A mixture of lots of different recordings so that in the studio there were microphones hanging above the audience in different positions and these were feeding into a multi-channel digital tape. And what we could do, would be to look at which microphone feed, the signal of the coughing was strongest in, loudest in, and by that process narrow it down to a particular area of the audience. And in fact, we narrowed it down to a set of people who were going to come on as contestants next – to four people, I think, or five, I don’t remember which. All except for one were men, so the woman could be excluded immediately – I mean it was clearly male coughing which left us with let’s say three people.

The issue was that next to the coughing, just on one occasion, the person coughed, and it looked like the contestant was going to give that as the answer, so you hear: “No!” It was the particular pronunciation of the word “no”, which is a monophthong vowel there, not a diphthong as in standard English, and there was only one of those people who would be a candidate.

It was a speaker of Welsh English, a Welsh accent, which has that particular pronunciation. So we went from a position of thinking it could be any one of 250 people, we’re never going to be able to sort this out to actually coming up with the cougher most likely to be this one person in the audience. And I guess that was a very high-profile case here, but it certainly wasn’t of the seriousness and the magnitude of something like genocides or murders or anything like that.

I think it had the nation amused rather than appalled that this person would be involved in what was essentially a very schoolboy-like ruse or prank to do. It was very simple. That was one case where we thought we're not going to get anywhere, and we did.

Another one I was involved in was a more serious issue. In the late 1970s, early 1980s in the UK there was a series of very vicious murders. Usually women in northern towns in the county of Yorkshire, and in Lancashire to a certain extent, and the murders involved hammer blows to the head of these women, who were on the streets at night, or on wasteland, and also a certain amount of mutilation of the bodies. Rather reminiscent of the Jack the Ripper murders in Victorian England – these were murders in the East End of London. And this person became known as the Ripper, the murderer. It took a long, long time to find him. And one of the reasons it took a long time for the police to find him was because someone, a hoaxer, had sent some letters and also a recording on a cassette tape through the postal system to the investigating officers to throw them off the track. The letters were all postmarked from a town called Sunderland in a county further north, so not from Yorkshire but from county Durham. And also the recording was in the accent from that area of the country and a colleague of mine, in fact my mentor, who has now passed away, Stanley Ellis, examined the recording. He had worked on the survey of English dialects, and he was able to pin this recording down to a particular place near this town of Sunderland or only two or three miles away.

And that threw the police off the track because the actual Ripper was a Yorkshireman. The police were convinced that the letters confessing to it and the cassette recording were genuine and therefore they had been looking for someone from that area of the country (Sunderland), not from the area of Yorkshire where the Ripper actually lived.

The Ripper of course was prosecuted and sentenced to life imprisonment when he was found. But for twenty-six years no one knew who the hoaxer was, no one had identified the voice, even though this recording had been played at the time of the murders at places like football stadiums, it had been played on television, on the radio. No one had said, "I recognize that voice it is...". I think it was in 2006, many years after, the police decided to open up an inquiry, a 'cold case' inquiry into who the hoaxer had been. And what they did was to take the envelopes which the cassette and the letters had been contained in all those years earlier, these having been stored by the police. In the days before self-adhesive envelopes, you had to lick the back of the flap of the envelope to activate the glue. They sent that away for DNA analysis and they got a cold hit on a guy from that particular area of Sunderland. It wasn't definitive. They took him into custody and interviewed him, they made him read aloud from the original transcript of the cassette recording and they also had natural flowing speech from him as well.

I was contacted and I was asked the question "Would you be able to compare the voice in the original cassette, the hoax tape, with the voice of the suspect?" I was given certain background information. But firstly, 26 years later there

are changes in the voice, natural changes which come with aging. But the background information I was given was that this man was a very heavy smoker, I can't remember how many – 20 or 30 cigarettes or more a day – and he also was drinking, very, very heavily drinking, lots of whisky every day and so on and so forth. Again, as in the millionaire case, the game show case, my initial reaction was “I don't think there's really very much hope. There will be changes in the voice owing to the alcohol, the cigarettes, and the passage of time will make it just the non-starter.” When we got the recording and compared it with the original cassette message, it was almost like the same person in the same room on the same day, and we were able to produce quite a strong report in that case about the likelihood of it being the same person.

Bijsterveld:

How do you deal with the fact that this cassette is very old technology when compared to what you would use today? Or would they actually make the recording with a similar type of equipment?

French:

No, they wouldn't actually have known anything about the machine on which it was made originally, it was never found. Nobody knows what the make or the model of the tape recorder was. I can't remember whether his police interview in 2006 was recorded digitally or onto cassette tape. That might sound rather strange, why would they be doing it on cassette tape? But the fact is that in certain parts of the country the police forces had bought, years earlier, mountains of cassette tapes for recording interviews and there were sheds of these things which had to be used up. I just can't remember whether it was recorded on a cassette tape or whether it was recorded digitally. But, anyway, it didn't hinder the investigation.

Bijsterveld:

Would you take these different recording situations and techniques into account in your report?

French:

We might not go into it in the report, but we'd certainly take it into account in conducting the analysis and it would be there in the records of analysis. It was interesting, there was another thing that we found in that case and that was that the original cassette recording was not continuous, that at various points, what he'd done was he recorded a section and then he would stop the recorder, he would listen to the section and if he was satisfied with how it sounded, he would record the next section. He was obviously reading aloud from a transcript that he'd prepared and then, in some sections, he'd go back over them, because he wasn't happy indeed and he re-record them.

But we could tell this using a particular technique. Firstly, the straightforward one was looking at the recording waveforms for any breaks in them. But the second one involved a particular technique – and you expressed an interest in collaboration with colleagues in East Europe and in Russia – it was a type of technology, which they essentially gave to us. It involved looking at the magnetic patterns in the oxide surface of the tape using ferro-garnet crystals.

These are a particular type of crystal which if you put up on the stage of a microscope and you look you can draw the tape over the crystals and the crystals take on the actual shape or the pattern of the magnetic fields in the oxide surface. That was useful. Being able to tell that the recording had not been made just all the way through, but it was obviously something which had been carefully planned, and it had been recorded in sections and he'd checked his performance, gone back, rerecorded bits, etc. What that went forward was as showing that this was not just some one of drunken caper, but it was something which showed method and system and premeditation.

Bijsterveld:

You said you learned to look at the gaps in that wave from colleagues in Russia, did you say that?

French:

Yes, that's correct. I mean, initially, we had a version of that sort of technology, which was called ferrofluid analysis. This involved using ferrofluid, which is a particular type of ferrous or iron particles which are suspended in a liquid carrier. We would apply these to the surface of the tape using a syringe. But in fact, the patterning was never as good as it was with this Russian technology, with these ferro-garnet crystals. And also, it was destructive to a certain extent of the tape. So you were at least partially destroying the original evidence, which made it difficult if another expert wanted to examine it. They could look at the photographs you'd taken, but they couldn't reproduce the test themselves. But the ferro-garnet crystals were once developed in Russia.

Bijsterveld:

When exactly did you learn from these Russian colleagues, when was it? Do you remember that?

French:

Well, we acquired the crystals during a visit to Moscow. I'm not sure how much I should say about this as it is going into a transcript for publication, but we, myself and another member of my firm, went to Moscow as an official UK delegation. I don't think we knew we were an official UK delegation until we got there. We thought we were just personal invitees, but when we got there, it turned out we were an official delegation. We spent some time with a senior police officer who was also a scientist, an engineer and scientist, visiting the labs in Moscow. Two labs. And talking to them about the sort of methodologies they used in authentication cases or speaker comparison, and a range of other things. And we came away from the visit with the crystals that we needed to do this sort of work.

At that point we were simply putting the crystal on the stage of a microscope and putting the tape over it. The crystals look just like parts of a mirror, you can imagine strips of glass which look like a mirror. What we later got was, again from Russia, a standalone system whereby these crystals had been embedded in a sort of platform, which had in its own lighting system, etc. But that technology, essentially, came from Russia.

Bijsterveld:

And when was this, this visit?

French:

I am trying to think about the exact date. I would say that visit was in the late 1990s, mid to late 1990s, I could look up the exact date.

Bijsterveld:

It is interesting. You also mentioned your firm. Can you say something about the firm?

French:

Yes, of course. Originally there wasn't a firm, there was just me. I was working as an academic in linguistics and phonetics. Around about the mid 1980s, I began to get consulted, just occasionally, in cases, which involved speaker comparison. And I was introduced to this sort of work by the person I mentioned earlier, Stanley Ellis, who worked on the survey of English dialects at the University of Leeds. At that time, there were only two people in the UK, well maybe three people in the UK, who were doing this sort of work. Very few cases a year, maybe just twenty cases a year in the whole of the UK or something like that. Stanley introduced me to the work, and I initially acted as a sort of understudy to him, a second opinion. My name didn't appear on the reports, I assisted him. It was very much in a sort of secondary capacity. Over a little period of time, I began to offer my own reports, maybe four, five, six a year, that sort of thing.

And then, in the UK, what opened the gate for the firm to be established was that there was a particular piece of legislation passed, which went onto the statute books in 1984. It was known as the Police and Criminal Evidence Act or PACE, as it's often given the acronym. It included a lot of things, like restrictions on how arrests were made, all sorts of things like that. But among all this, was a decree that all interviews with suspects at police stations should be recorded, at that time tape recorded. Prior to that, if somebody had committed a voice crime in the UK, let's say they'd left a death threat on somebody's answerphone or whatever, they would be invited to go to the police station and when there, they would be invited to give a voluntary voice sample. Of course, most people would say "well, thank you very much for the kind invitation, but as it's voluntary I will decline." So there was nothing to compare the criminal recording with. But, of course, with PACE, this new legislation meant then all of a sudden, in every case, there was a comparison sample of speech from the suspect, in other words, the recorded interview.

So the interviewees found themselves inadvertently, almost as a by-product of the interview process, providing a voice sample. And the cases went from maybe twenty-five a year in the UK to a very, very large number. Nowhere near as many as there are nowadays, because people weren't recording so many things then.

That opened the gate to establishing the firm of mine. I came out of academic work for a while and just did speaker comparison and other types of forensic speech and audio analysis, initially working on my own with maybe a secretary who came in one morning a week to type up the reports and send out bills. And then it became clear this was going to be a sustainable career and that

I should be working with other people and taking independent premises. So sometime, in the late 1980s, I can't remember the exact year - maybe 1988, 89 or 87, I established a firm which takes my own name *J P French Associates*, and it's been established since then. And it's been variable in size. I think at one time we had maybe eight or so employees, at the moment we've got five. And we do the largest share of prosecution work and defence work in the UK, and we've been involved in cases right across the globe: the Southern Hemisphere, Far East Asia, Africa, lots of European countries.

I guess that's the history of it. The UK is rather different from, say, the Netherlands, or Germany, where they have this sort of work done in police or state laboratories: so you [the Dutch] have the NFI, the Germans have the Bundeskriminalamt [BKA], which have their own speaker comparisons, the voice identification division, and the various LKAs, Landeskriminalamten, have similar labs. In the UK that's never happened, so all of the voice comparison work goes out to private tender. To either a small number of university academics, who are prepared to do it, or to people working privately or to a lab such as ours.

Bijsterveld:

So you work for both parties then?

French:

We work for both parties, the prosecution and the defence. However, I would say that historically, our work has been seventy-five maybe eighty percent prosecution rather than defence. But that's simply because there's more prosecution work around than there is defence work. It's not because we have a preference for working for the prosecution, but in a lot of cases where the prosecution brings this sort of evidence, the defence doesn't contest it, therefore there isn't defence work to be done. But we work equally for either side, whoever gets to us first, and we've always taken a pride in the fact that our work is impartial, and I'd like to think objective, that we don't see ourselves as being in any way tied to the people who are enlisting us or who are paying the bill. It's simply a function of the English or the UK adversarial system of justice, that one side or the other has to approach us, and it is very rare indeed to be appointed by the court as it is in the Netherlands or Germany, et cetera. And we'd like to think that in every case the report we write would be the same report if we'd be enlisted by the other side.

Kvicalova:

I have one additional question about the experts - is it common that more than one expert is invited to write the expert review, or is it only in the cases when the defence wants to have its own expert? Is it a standard practice that more than one expert review is written?

French:

No, there would only be more than one expert if the defence wanted to contest the prosecution, or in rare cases the defence might have expertise of this kind where the prosecution hasn't, and the prosecution would want to contest the defence. Where there's an expert for each side it's in cases where the evidence is in dispute. However, in another sense, there are two experts in every case.

In the UK we have a quality control and management system, whereby each expert, say within my firm, who produces the main analysis, they would be the primary expert. But their work would always be checked by a second expert. And this is going to become law in the UK shortly – I will talk about that in a minute – but the way that would work would be the first expert would do the full analysis and they would come to their provisional conclusion about the issue of the voice in the two recordings. When the first expert had completed the work, the second expert would be brought in, and they would do two separate tasks. One would be to check all the key findings, the critical findings of the primary expert, so they would be given those findings and they would be given the recordings and they would go through, and they would check the key findings against the recordings to see if they agreed with them. The second bit would be what would they make of the findings if it were their case. So, they wouldn't be shown the conclusion of the first expert, that would be kept from them. It would only be after they'd come to a decision as to what their conclusion would be that they would be shown the first expert's conclusion.

Usually these would be the same, in some cases they might be slightly different, and we don't really have cases where they're a long way apart. But if they were slightly different, they would have to discuss the case and reconcile the differences and decide which would be the conclusion to go with. So, in that sense, there are two experts involved in every case, but that's experts working on the same side in the same lab. Not all UK cases do have that system, though, because you need a colleague to be working with you in order to do it, and some of the UK experts are sole practitioners so it's not possible. And really, the way we work in the firm is that we have three people involved in each case. So, the primary expert, which I've talked about, the secondary expert who is the checker, but we also have a third person who is an information manager. In order to reduce the possibility of cognitive bias, the third expert will filter the material in such a way to keep out things that could be seen as very incriminating. So, let's say it's an interview recording that's going to be compared with the criminal recording. In the interview recording, there might be material whereby the suspect incriminates themselves. So, the information manager would be responsible for editing the recordings before they go to the first expert so as to exclude as far as possible – it is not always possible to exclude all – but to exclude as far as possible material of that kind, and also to shield the first expert against information, which comes directly from the instructing party. That's a police officer or a defence lawyer. This is things like “Well, to me it's perfectly obvious it's the same voice in both recordings, this man's a known criminal, he's done this sort of thing before”. That's exactly what the first expert, as well as the second expert, doesn't want to know about. So that information is put aside. In that sense, I guess, there are three experts involved each case.

Bijsterveld:

That is super interesting.

French:

We now have this position, which is a government position, a Home Office position, called the Forensic Science Regulator. And the Forensic Science

Regulator's role is to put all areas of forensic science, whether that's DNA, psychology, blood or forensic speech and audio analysis, under government regulation. This involves writing a standard to which people have to work, and also making sure that people put in place quality management systems of the type that I've talked about, and a whole range of other things. At the moment, the regulator now has statutory powers, and an initial version of a document towards establishing standards is about to go into parliament.

It looks as though it's going to be about four years down the line before the forensic speech and audio specialism is brought fully under statutory control. But it's there, it's on the books, it's imminent. And my firm has moved towards the standard A) because we think it's a good thing and B) in anticipation of the statute being enacted.

Bijsterveld:

You already said something about how you work normally. Obviously, you prepared the questions very well. What would you like to say about the second question, that is, what does your work routine look like in practice? I know it is a difficult question, because it is a rather comprehensive question. In terms of the technology you use, the procedure, the cooperation, the workspace, et cetera? Perhaps the relation between what you do, the machine part and the listening part? We are really interested in that.

French:

We use Dutch software! We used to say it involved two parts. The other people, Herbert [Masthoff]¹¹ and Angelika [Braun], have been through this. The auditory analysis and the acoustic analysis. And the way we used to present these was that these were separate things. In reality, these things go ahead in parallel. So, one is listening and at the same time taking acoustic measurements from specialist software. And the system which we use, the software system is very very well known, and I think it has probably become industry standard; it's called *Praat*. It is freeware now, produced by Paul Boersma, who's a Dutch academic. He's made it available, and it's more or less universally accepted both in university phonetics departments and in the forensic context. It's made the work very much easier than it was. When I first started in this area, if you wanted to do acoustic analysis, you had to buy really big expensive dedicated standalone instruments, which was hard. Now, of course, you can do it on any PC using the software alone.

So, the way we work is we re-record the material within the software, which allows us to listen to it short section by short section doing auditory phonetic notations. We have a specialized system of symbols, the International Phonetic Alphabet, which we use to capture the fine-grain detail of how consonant sounds are pronounced and vowel sounds are pronounced. And we make a record of that, and we compare whether they're pronounced in the same way across the two recordings, using our listening skills developed as part of a university education in phonetics. So, we're not listening so much to what people are saying but how they're saying it. How the consonants and vowels are pronounced. As part of those auditory investigations, we would be making

¹¹One of the participants in the witness seminar on speaker identification, 17 June 2022, see above.

a note of patterns of intonation. In other words, the melodies of the speech. If there was anything there of interest, we would be noting that. We'd also be looking at speech rhythms. In certain cases, we would be looking at rate of delivery, we'd be averaging the number of syllables per second. If we thought somebody is a particularly fast speaker or particularly slow speaker, we'd bring that in. And we'd average that across whole series of utterances. We also look at things as part of the auditory and the acoustic testing such as disfluencies, that is *uhms*, and *ehrs*. Whether the people when they're their pausing, they make those sorts of sounds or whether they just produce a short silence, or whether they, let's say, string out consonants and vowels. So, we have a scheme, which we use for categorizing what their disfluencies are and we can look at the disposition, the proportions of what type of disfluency versus another in each of the recordings.

At the same time as we're doing this, and it really is in parallel, it's not separate examination, we are looking at displays on the computer screen of the speech. In other words, we make extensive use of spectrograms. We make use of them in two ways. Firstly, is there acoustic corroboration for what we're hearing. So, it is used as a corroborative thing - do we see the same thing on the screen that we're hearing. That's put very simply.

The other use that we make of it is there are certain things which you can detect from acoustic analysis that you can't from simply listening, and vice versa. So, let's say, some of the higher formant frequencies - these are resonances usually within vowel sounds, which are at high frequencies - you can see those, you can resolve them, you can measure them on the computer screen, but I certainly can't hear that sort of information without examining the speech acoustically. So, as well as the acoustic analysis corroborating - or parts of it corroborating - the auditory analysis, the listening analysis, it also throws up supplementary information, or complementary information, which wouldn't be available from listening.

And what does it look like? Well, the workspace looks very much like the sort of office you might find in an insurance firm or anywhere else. You know, twenty, thirty years ago, it looked like a stage set from Star Trek. Lots of big instruments with dials and flushing lights, and knobs and now it's just, generally speaking, PCs on desktops with the software on them and a pair of headphones attached. It looks less exciting, the workspace, than it used to. People, TV companies, who visited us at said times have been rather disappointed to see that, and expected something much more hi-tech looking.

Bijsterveld:

To go back to this Star Trek episode: you said that you had to bring large instruments. Do you still have information on what you exactly used, or do you have photos? Because for us as historians, this is of course interesting.

French:

When I first began to use this technique of spectrography, which is to visualize sound energy across frequencies over time, I had an instrument produced by a firm called *Kay Elemetrics*, and it was called the sonograph. It was about two

and a half feet square by about a couple of feet high, and on the top was the revolving drum. And onto this drum, you fixed a certain type of paper and there was a stylus, which, I think, carried an electrical current and that used to get hot, and it would etch into the paper the patterns which we now see on the computer screen in spectrograms. It had an exhaust pipe which consisted of one of those flexible hoses you might see on hoovers or on the back of tumble dryer machines. And this went from the machine out of the window into the street because we were told that the fumes that came from the etching, the burning of the spectrogram, were carcinogenic. The general position was to send them out into the atmosphere, Not a very good solution but better than into a small enclosed room. That was the instrument.

We also had another standalone instrument for measuring pitch, it was called a pitch meter. And we had something called a cepstral analyser. I won't go to the technicalities of it, but some of this technology, in the very early days, it was very difficult to get hold of it. Some of it came under government control, it was held as an official secret, if you like, and it was only released to certain government agencies, and for people working in the private sector it was difficult to acquire. That was gradually loosened.

Bijsterveld:

When was that? Was that also after 1989, or when did it gradually loosen?

French:

It depends on which instrument we're talking about. The technique of spectrography was developed in the 1940s under conditions of very strict secrecy. It was developed by Bell labs in the USA. They only came public with it in 1945 when the soldiers were coming back from war, a lot of them with hearing damage, and they said they developed it so that people who were deaf, the war veterans, would be able to visualize speech. The reality was that it hadn't been developed for that at all, it was developed as a technique to help with encrypting speech. When there were radio contacts between British installations and ships and things like that, they could be encrypted, using a version of this technique, which scrambled the signal. It took a long time for that to be made known, and the official story was it had been developed to help these deaf war veterans.

Bijsterveld:

So you are referring to the Potter report regarding visible speech, right?¹²

French:

Yeah, I mean at this point there was quite a number of documents relevant, actually. But spectrography as a technique was not secret, but there were certain things, like types of cepstral analyzer, which were not publicly available even into the mid 1980s. And there were also certain types of sound filter, which had been developed largely in the States, where the technology was available to security services, people working in that area -sound filters being devices which are designed to suppress background noise in recordings, make speech more intelligible, etc. These were in the sort of grey area between public domain and

¹² Potter, R. K, Kopp, G. A. & Green, H. C. (1947). *Visible Speech*. New York: Van Nostrand.

not quite public domain - which it was possible to acquire, but not the whole range of them, and that was certainly so in the mid 1980s as well.

Bijsterveld:

So how did you get that equipment after all, what did you do to get it if it was difficult to get?

French:

We had to wait until these companies considered they could sell it to the general public or whoever the governments allowed them to. And, of course, there were certain companies who would say: "Well you know, we couldn't possibly sell you this, we're under government control as to whom we can sell it to and it isn't to go out to the private sector. Or if it did, it would be very expensive". So, in other words, there were certain companies who might in return for quite a large sum of money release it. And I'm not going to name them.

Bijsterveld:

Would this define what kind of equipment you would buy, or could you still make a comparison of what worked best?

French:

At that time, I think it's fair to say that in the UK people would not. Until we began to do the acoustic analysis, it's difficult to put a date on it, but I would say in the early 1990s or so, before we were doing acoustic analysis in conjunction with the auditory analysis, and really the people who adopted acoustic analysis took what was available or what they could afford. Certainly, by the mid 1990s there were PC and Apple Mac programs for doing spectrography. Certainly, the Apple Mac technology was still very expensive and gradually it came down in price as did the PC, up to the point now where it is freeware, as I mentioned before.

Bijsterveld:

Has that changed the field?

French:

It has changed the field a lot. I mean, people, unless they were actually either working within a university or they were involved in this as a dedicated profession, they wouldn't have been able to afford the early technology.

Kvicalova:

I would like to come back to writing the reports and ask you about communication strategies. How do you communicate the results of your analysis to non-experts and, in particular, how much are sound spectrograms, sound visualizations, useful in the process of communicating the results to the judges, lawyers, police officers, to the non-experts?

French:

The way we write our reports at the moment is that we write, at the beginning, a very non-technical summary. You know, the sort of summary which our colleagues in universities might wince at. It cuts corners, it talks about it in plain English. And then another part of the report sets it out in a bit more detail

and if people want more and more detail than they would ask for the records of analysis.

The reports are somewhere between being very technical and being accessible. Certainly, the summary is accessible to anyone.

What you've identified, Anna, is I think one of the most significant problems that we face, and that is how to communicate the results to people who are non-experts, who would say, "Well, come on what's the result, is it him or is it not him?" We would never express the result in that form anyway. We used to, you know, in early 1980s – that's exactly how we expressed it: "It is my considered professional opinion that the voice in the two recordings is that of the same person", which was a very extravagant and, looking back on it, I would say unjustified and dangerous sort of claim to make. Now the conclusion wouldn't be set out in that way. What we would be doing would be expressing it as the verbal equivalent of a likelihood ratio. You know, "the evidence we found in this case is more likely on the view that the suspect and the criminal are the same person" or "it's more likely on the view that the suspect and the criminal are different people." We wouldn't even put that in the report, we would move to what's usually known in forensic science as a support statement. In other words, the conclusion would be something like: "The evidence we have found provides strong support for the defence view of their being different people" or "the evidence we've found provides moderately strong support for the prosecution view that these are recordings of the same person". We work with a scale of support which has been adopted by the Association of Forensic Providers in the UK and it's common across all forensic disciplines. So, we don't ever produce categorical conclusions of it being the same person or it being definitely different people. We use this scale whereas in earlier days, first of all categorical "my considered view it is the same person or not the same person", then we moved to a classical probability scale, which had in things like "on the basis of the evidence I found the suspect is probably the defendant" or "very probably the defendant". That was a fixed scale but now we work with the likelihood ratio approach, and we express the conclusion in the report only as a support statement. Whether it supports the prosecution view or the defence view and to what degree.

Kvicalova:

May I ask you when approximately was this likelihood ratio and the scale introduced in the UK?

French:

I think the scale that we use now, the support scale my firm adopted, I think it was in 2014, so eight years ago. Before that we had something rather similar to it, but which was not very useful. It was an interim thing, it's often referred to as the UK Position Statement. It was something which was more or less just used in the UK, which we devised, but it was contentious. So, we came over to this, which is a much more useful way of expressing the conclusion, a much fairer way.

Kvicalova:

The classical probability scale “with high probability he is guilty, with the highest probability”, et cetera – was it used already in the 1980s, or not at all?

French:

That scale developed in the 1980s and there were various versions of it around and before that. I don't think I was ever involved in doing categorical identifications, but certainly my predecessors, Stanley Ellis of the University of Leeds, John Baldwin at UCL in London, maybe one more, they were involved in categorical identifications.

At the point where I got involved, I was very uneasy about categorical identifications. I think the classical probability scale is something that I was initially responsible for in the UK and that was disseminated from me to other experts working in the area. And then over time we went to this sort of transitional thing called the UK Position Statement which, as I say, wasn't very useful.

And we published that in the *International Journal of Speech Language and the Law*, that attracted various criticisms that it wasn't a proper Bayesian approach, it wasn't likelihood ratio based, and I think in 2014, we went over to the current approach, which is Bayesian, in the sense that would consider the likelihood of having got the evidence that we have on the basis of two different propositions, the defence one and the prosecution one. And there's also a verbal likelihood ratio. We can't produce numerical ones because we don't have the population statistics. But there's no reason for not doing it verbally and certainly the support scale is something which is very easily comprehended by judges and juries, you know.

If you wanted copies of the various scales that we've used across the years [...] There is an article I wrote in 2017, which is quite a long article, it's called “A Developmental History of Forensic Speaker Comparison in the UK”.¹³ You might not actually find it. I published it in a journal, which I assumed was going to be more or less internationally available. It's a journal produced in Japan, but it's notoriously hard to find the presence of this journal on the internet. But I will send you a copy.

Bijsterveld:

Wonderful, thank you. The 2011 overview with Erica Gold is easy to find.¹⁴

French:

That was superseded by another one I did with Erica Gold in 2019.¹⁵ There are some important differences over time, so if you look at the 2011 one, you will find that we haven't got onto automatic speaker recognition systems, but at

¹³ French, P. (2017). A Developmental History of Forensic Speaker Comparison in the UK. *English Phonetics*, 21, 271–286.

¹⁴ Gold, E. & French, P. (2011). International practices in forensic speaker comparison. *International Journal of Speech, Language and the Law*, 18(2), 293–307.

¹⁵ Gold, E. & French, P. (2019). *International practices in forensic speaker comparisons: second survey*. *International Journal of Speech, Language and the Law*, 26(1), 1–20.

that time there was only something like seventeen percent of the labs and practitioners that we surveyed internationally who were using an automatic speaker recognition system. In 2019, again, you need to look at the article or I would need to look at it, but I think the figure had moved up to something like forty-seven percent.

So, there were some major changes in that period, but still I would add now that no lab that we surveyed in either 2011 or 2019 was using automatic speaker recognition software as an alternative the auditory- acoustic approach. It was being used as another tool in the toolbox for the speaker comparison expert and it was always accompanied by some sort of human listening and/or acoustic analysis as well.

Bijsterveld:

We quite naturally move to the next question about which events have been crucial for how your field developed. You already talked about the ways of expressing evidence in courts, and the skills you use, and the move towards using the machine analysis, which was never used as the only way of analysing. Is there anything else that you would like to flag?

French:

It is not a landmark thing, but it's something you should understand in the difference between the UK and other countries and that is the use of the automatic speaker recognition system. My own view is that they're an extremely useful tool, and that they should be incorporated into the analysis. I would say that would also be the position of probably the majority of experts internationally. Nobody that I know would say, perhaps one person only would say, that they should be the only type of analysis. I mean, maybe one person that I know, but everybody else would say "yes they're a useful tool". Well, some people actually don't say it is a very useful tool, they want to keep them out of it altogether, but of those people who do consider them being useful, they would say "yes, incorporate them, but don't give them a massive special status, it is another tool. Sound spectrography is a tool, measuring voice pitch is a tool, and the use of an automatic system as a tool." So, it would be a tool among many tools that gives them the analysis. And that's my own position, I think it should be brought in and particularly in the light of the error rates that these systems produce that come down and down and down and down, where with good quality material, they produce error rates like two or three percent, very low error rates. So, I think there's a very, very good case for incorporating them.

I did some research with a colleague, and we found, and this is ten years ago or even more, we found out that in about seventy-five to eighty percent of forensic cases, the recordings were unsuitable for processing with an automatic system, they were either too noisy or too short. But the technology has got better, I would say that's going down to maybe fifty percent now rather than seventy-five to eighty percent. But in those cases where it can be used, I think it wouldn't be useful just to dismiss it out of hand. In the cases where it can be used it should be used, but in association with the other method, the traditional method.

But the UK has a problem, the Court of Appeal has said that in the present state of knowledge, an automatic speaker recognition system should not be used. The UK is a common law system, it is not a civil law system like most of continental Europe. In a common law system, law comes from two sources, it comes from statutes produced by parliament, but it also comes from the superior courts. So, if the superior courts pass a ruling of that sort, that a particular forensic technique should not be used, then for lower courts, the ordinary criminal courts, it's binding upon them not to use that particular technique. And we have a UK ruling from the Court of Criminal Appeal, which - I told you what it says. So, we cannot use automatic systems in the UK. They are routinely used in Germany, in the BKA, and I think they are in Holland, certainly the NFI were experimenting with them, it could be, I would have to check the casework. They are now used routinely in Spain, they are used in a whole range of other countries, but the UK has got this particular ruling, which more or less says "no, they're out of court".

Bijsterveld:

You said that it is never used as the only method. That is also how I understood it from Maartje [Schreuder]¹⁶, for instance, but you can bring the evidence to court in the Netherlands. And you are saying that in the UK you cannot do anything with it? Because you said it is important to use it. So how does it relate to what happens in court?

French:

Well, it doesn't in the UK. I think you should be able to use it, but we can't in the UK.

Bijsterveld:

Okay, so it is simply not part of your reports, then?

French:

No, I think it should be. It is used in the UK for intelligence purposes, but not in production of evidence in court. Law enforcement agencies can use it for identifying voices of suspects, etc., but nobody can use it at the present time for evidential purposes.

Bijsterveld:

That is interesting.

French:

I'm involved in various negotiations at the moment about ways of circumventing that particular ruling and bringing it into play in forensic casework.

Bijsterveld:

And when was this ruling?

French:

Either in 2017 or 2018.

I think it was in response to my attempt to introduce this sort of evidence to an Appeal Court case. I obviously wasn't convincing enough as a witness because this was the outcome.

¹⁶ One of the participants in the witness seminar on speaker identification, 17 June 2022, see above.

And interestingly, as you said, it's always used in conjunction with the traditional method. One of the bases on which the Appeal Court founded its ruling was that "if this guy [i.e., French] believed in this automatic recognition system sufficiently, why would he be using it in conjunction with the traditional method?"

Kvicalova:

May I ask about that? What I get from the discussions we had in June¹⁷, and with other colleagues who work, for instance, on the automated analysis in US, is that part of the criticism of the automated approach comes from the fact that, sometimes, it is not actually used in conjunction with phonetic, aural, and acoustic methods of analysis. That there are even companies and experts who just specialize in only this one particular type of expertise. So, who should be the expert to perform the automated speech analysis? Do you believe that it should be a phonetician or somebody who has a broader conception of how the voice actually works, and uses this type of approach only as a part of the overall analysis, or is it okay if there are experts who just specialize in automatic machine recognition?

French:

No, the former, the first thing you said. It should be done by someone with a training in phonetics. With a training in speech science, phonetics, who knows how the voice works, who knows what you should be listening for, what you should be analysing acoustically and also with some conception of Speaker variability. Nobody knows exactly what the automatic systems measure, they are a little bit of a black box, but we have a very good idea of the dimensions of the voice which they measure. In other words, it would be used by somebody who knows what they're doing.

Kvicalova:

But it is not always the case. Even in the Czech Republic, there is one expert who specializes in automated analysis, there is another one who is a phonetician, and a third one who is also a phonetician by training. So, it is probably not everywhere as you describe it should be, or as your ideal vision for the UK is.

French:

If they are people who are not speech scientists, who are using the system, I would be very suspicious, maybe circumspect about using those experts.

BREAK

Bijsterveld:

It is interesting to see how the field itself looks back and how different this is for different countries.

French:

I find the history of the development of the science and the events surrounding that and events that actually trigger off changes in perspective, I find that immensely interesting. In fact, in addition to the developmental history article

¹⁷ During the witness seminar on speaker identification, 17 June 2022, see above.

that I wrote, I am currently writing another one for a big textbook which is on legal aspects of it, like how the evidence was treated in the UK initially, when the first case was. The real focus is how it was received in the courts. Are there any Appeal Court rulings or Superior Court rulings, which bear on it? What dates did they come out? For instance, in England we have different jurisdictions here. We have three: England and Wales together in the UK are one jurisdiction, Scotland is another, Northern Ireland is the third. In England and Wales, the issue of whether auditory analysis should always be accompanied by acoustic analysis for the evidence to go into court, that was considered by the England and Wales Appeal Court in the 1980s and they decided “We’re not going to rule on this, it’s up to the expert”. The Northern Ireland Appeal Court had the same decision to make with respect to another case, and they said “if it’s prosecution evidence there must always be acoustic evidence, and the particular type of acoustic evidence will accompany the auditory evidence”.

More recently, the England and Wales Court of Appeal had to consider the same issue again, and they said, “we will not go as far as the Northern Ireland Appeal Court, we’ll allow it in if it’s auditory evidence only”.

I guess there are things which people don’t realize about what these rulings mean. Obviously, they are binding on the lower courts in the UK, but also, in the absence of other more local rulings, they are not strictly binding, but they’re likely to be highly persuasive to courts in other Commonwealth countries. I don’t remember how many Commonwealth countries we have now. I always think it’s just the UK, it’s Canada, it’s New Zealand, it’s Australia, but in fact there is something like fifty of them. So, these rulings that emanate from London or from Belfast or wherever, actually have an effect on the way evidence is handled in these other countries. So, at the moment what I’ve been doing is documenting all this sort of stuff about the evolution of these common law Appeal Court rulings and all their influences across the globe.

I think your field is actually probably more interesting than mine.

Bijsterveld:

Who is editing the book you have mentioned?

French:

Kirsty McDougall from the University of Cambridge with two of her colleagues, Toby Hudson, and Francis Nolan. It’s Oxford University Press and it’s called the *Oxford Handbook of Forensic Phonetics*.

Bijsterveld:

That is super interesting, because one of our authors [from the workshop *Forensic Voices*], Michael Mopas, is also writing on how speaker identification expertise was actually received in US courts.

French:

I’m trying to take a broader international perspective. For instance, in the USA, they’ve got a number of rulings which are to do with the admissibility of expert testimony. There’s one of the Federal Rules of Evidence, which has a bearing on it.

They're also Superior Court rulings, two in particular. The best known one is Daubert standard which everyone talks about.

I am looking at the position in the USA versus the UK. But I also surveyed probably about fifteen people, forensic speech scientists from different countries, including Czechoslovakia, and lots of other countries, and asked them what is the position there? What are the rules of admissibility on scientific evidence and forensic speech evidence in particular? I've got quite a lot of data. In fact, it extends to places like India and Chile.

Bijsterveld:

We are looking forward to reading it!

There are two remaining questions we would like to ask you. We already talked a lot about the opinions regarding the best and the most informative way to do speaker identification and how that changed over time. We have already covered that to a large extent. Is there anything you would like to add to what you have said so far on that?

French:

Nothing in particular, no. I mean, initially in the UK, the auditory approach only, then auditory-acoustic, then an attempt to bring in auditory-acoustic incorporating automatic, which failed. I think that's about it, really.

Interestingly, if you look at the USA, it was the other way around, their first attempts to introduce this, when we were doing auditory only, they were doing acoustic only, the so-called *voiceprint* tradition. They would simply look at spectrograms of the suspect saying the same words and they wouldn't even take measurements, there would just be a sort of crude form of picture matching "do these match or are they different", but they are always different. So that was the sort of converse position there.

But I was just thinking during the break about the thing that you mentioned about people doing ASR [automated speaker recognition] only, who are not speech scientists. One of the famous cases in the USA where the whole prosecution case came off the rails was a guy who was doing just that, a non-speech scientist. That was the George Zimmerman trial, which I think Angelika [Braun] has mentioned and I can send you a link to the evidence in that. It was all video recorded and it's all available on the Internet. And that shows what can happen if somebody who isn't a speech scientist just operates an ASR. They can get it horribly wrong, in ways which mean the evidence can't go in court.

Bijsterveld:

Our last question is to what extent you had contacts with speaker identification experts in Eastern European countries, especially pre-1989. I don't know how active you were back then already. And also, what do you know about differences and similarities in how the work was done at both sides of the Iron Curtain?

French:

I don't really know what happened before the Fall of the Wall, my contacts with the Russians, principally, were all in the 1990s, probably mid 1990s to late 1990s. At that time, the lab that I went to, which is the main lab in Russia, the way they

were operating was that they assigned two analysts to each case. One analyst was an engineer, and the other analyst was a phonetician/linguist.

And the details of what each of those people did I either never really got, or if I did, I have forgotten them. There was some traditional phonetic analysis going on in there, an assessment by a linguist, and also, I think, the engineer undertook the acoustic analysis, but again, I can't really remember what that involved in detail.

My impression was – and if ever my hosts from that time listen to this, I hope they're not offended – that they were rather hidebound, that they had an absolute fixed protocol, which they applied in every single case, irrespective of what the peculiarities might be of a case. They had this sort of framework imposed upon it, that was the overwhelming impression that I came away with.

There was a return visit, whereby the guy who was the head of the lab came to my lab in the UK, spent some time with us. And I think he was, on the one hand, quite taken with the idea there should be flexibility in approach, so that you could change what you look out for from case to case on the basis of what appears to be important from an initial listening to the recordings.

But, on the other hand, I think he found it rather strange that we didn't have this very rigid approach, a sort of grid through which we looked at each case. Which, I guess, would correspond to people's perceptions, stereotypes of how Russia was.

Bijsterveld:

It sounds very similar to how it was done in the GDR. Everything was focused on doing it as standardized as possible. Even though the political situation was full of bias. It is interesting to see these two sides of the coin of how the sciences worked in that context.

French:

There was also private work going on in Russia at the same time. In the International Association for Forensics Phonetics and Acoustics we had a member from Moscow State Linguistic University, her name is Rodmonga Potapova. She was doing casework in Russia at the same time as a private expert. I never really got into the details of her methodology with her, but she was a speech scientist, and she knew what she was doing.

Bijsterveld:

Was it your impression that the Moscow lab had been there for quite some time?

French:

Rodmonga's lab or the state lab?

Bijsterveld:

Well, both then.

French:

The state lab, just going from memory, I think was something like thirty, thirty-two people working in that lab. So, they'd obviously built up the staff,

they'd built up this protocol and these procedures. I don't know when it was established. If you wanted to know that out, the guy, who was in charge of the overall setup of forensics within that area was a guy called Dr K [anonymized]. I'm sure he would correspond - well, maybe under the present climate it would be more difficult, but he would have been quite open to approach.

Bijsterveld:

And the private lab?

French:

The private lab would be of Rodmonga Potapova. She started possibly in the 1970s, 1980s, I'm not sure. I think she'll have retired by now.

Bijsterveld:

Is there anything else you would like to share about these relations?

French:

In the 1990s probably, I'd have to look up the date, there was a special session set up in a conference called SPECOM, which was held in a hotel miles into the forest about 40 miles outside of Moscow. It was an international conference and there was a special session in that on forensics. I was invited to go there and give a plenary lecture, and I met there various people from the Russian government and police and security agencies, who were doing this sort of work. None of them, as I recall, had any background in speech science, they were simply using engineering techniques, automatic systems.

Bijsterveld:

What is interesting is that we found that in the GDR [German Democratic Republic] what they really were after was indeed something fully automatic. That's what they wanted. So, the auditory analysis is what they saw as "okay we still have to do it, but what we really want is a system in which we can do it by fully automatic means." So, it is interesting to see that, apparently, this was also the goal in Russia and also what they wanted.

French:

At some level, it is desirable that you could simply put two samples into a machine without any human intervention and get a readout. In the way that you do, say, in toxicology analysis, you know, twenty percent chalk, seventy-five percent diamorphine. It would be wonderful if we could do that, but for a whole range of different reasons that's never going to happen in speech comparison work.

While it's desirable because it maximizes objectivity, it takes out entirely the subjective element. Although from a judicial point of view it is fantastic, I think it is never ever going to happen. And there are very good reasons for that. People just vary from occasion to occasion in how they speak. Lots of factors involved: emotional factors, illness, smoking, drugs, tiredness, it's whole range of things like that. And also there's no empirical basis for assuming that everybody's voice is unique, there could well be people in the population, I don't know how many, but there could well be people in the population who are sort of the vocal equivalent of doppelgängers. That we just don't know.

So, I don't think we would ever achieve that.

Bijsterveld:

Why do you think that in Russia the engineer was so important?

French:

Do you mean in the labs, why there should be one engineer and one phonetician?

Bijsterveld:

Not really. You spoke about the representatives of the Russian government, the police, the security agencies, and you said they were very much into ASR. That's what I understood.

And it is also our task to better understand why that was the case, and we see that there has been a long tradition and that they were really after this approach. Because you talked to these people, perhaps you could sometimes pick up some of their views or visions or assumptions.

French:

I can't really say why that should have been the case in Russia. All I can say is that once a tradition starts that somebody produces some software that draws on a particular sort of engineering background, that grows but takes it on its own impetus. My impression was that with the exception of Potapova and maybe a couple of other people, who were not in government employment, there was nobody in the government sector who was a phonetician, a linguist, except the people working in the state lab, but they were, as I say, working in conjunction with the engineers. I don't think the engineers at that time were actually utilizing an ASR, or if they were, they were just beginning to. I think what the engineers were doing were the acoustic measurements, which their phoneticians wouldn't do. What the phoneticians were doing was listening to the recordings and making impressionistic judgments.

I mean the attractiveness of the ASR is, firstly, the reduction of the subjective element that we talked about, which is always there with the phonetic approach, but also the fact that using an ASR, you can process hundreds of samples in a matter of seconds, if not a matter of minutes. Whereas using the traditional approach, comparing one criminal sample with one suspect sample plus the writing of a report, you might be looking in the region of about fifteen hours. You can see that it is an economy in terms of time and cost.

Bijsterveld:

That actually matches quite nicely with what the GDR people were saying at the time, they talked about between fifteen and twenty hours.

Kvicalova:

This might have something to do with the changing notions of objectivity in the field. I was wondering whether you think that there were any important shifts in how objectivity was understood in the field. Whether, for example, there was ever a widespread notion that the ultimate exclusion of the human listener and its replacement with automated speech analysis would make the analysis really more objective, or whether nobody ever believed this.

French:

I, at some level, believe that if you were able to exclude the human listener, then it would make it more objective. But the problem is you're never going to be able to exclude the human listener. Or if you do, it would be more objective, but it will get the results wrong. It's too difficult.

Bijsterveld:

It is interesting what you said about the situation in Russia at the state lab—that indeed you might have what we would call path dependency. You start in a particular mode and then it develops more and more into that direction.

French:

I think things just take on their own momentum, once they will acquire a sort of critical mass, that mass gets bigger and bigger. It often happens in science, I think.

Bijsterveld:

That is right. At least that is one of the claims in our field that type of technological networks might get so much momentum that it is really difficult to undo it.

I think we have covered all the issues we wanted to discuss today, thank you very much for sharing your reflections with us!

04

BIOGRAPHICAL
NOTES

BIOGRAPHICAL NOTES

Editors:

Anna Kvicalova is a historian of science, religion, and the senses; in her work she deals with the history of sound-based knowledge and listening skills. She works as a permanent research fellow at the Centre for Theoretical Study (Charles University and the Czech Academy of Sciences) where she is the leader of the research project *The Second Sense: Sound, Hearing and Nature in Czech Modernity*. She received her M.A. from the University of Amsterdam and a Ph.D. from Freie Universität Berlin. Between 2013 and 2017 she worked as a research fellow at the Max Planck Institute for the History of Science in Berlin. She is the author of *Listening and Knowledge in Reformation Europe* (Palgrave, 2019) and other publications on sound, hearing, and acoustics (published in *Technology and Culture*; *Annals of Science*; *Sixteenth-Century Journal*).

Karin Bijsterveld is full professor of Science, Technology and Modern Culture at Maastricht University. Her work focuses on the cultural history of sound in its relation with science and technology. She is author of *Mechanical Sound: Technology, Culture, and Public Problems of Noise in the Twentieth Century* (MIT, 2008), co-editor of *The Oxford Handbook of Sound Studies* (Oxford UP, 2012, with Trevor Pinch), co-author of *Sound and Safe: A History of Listening behind the Wheel* (Oxford UP, 2014, with Eefje Cleophas, Stefan Krebs and Gijs Mom), and editor of a special issue on Auditory History for *The Public Historian* (2015). Among her other books are *Sonic Skills: Listening for Knowledge in Science, Medicine and Engineering* (Palgrave, 2019, OA), and the edited volume *Uitgepakt/Unboxed* (Verloren, 2021, OA) with popularizing entries on scientific instruments. She is currently working on an edited volume on *Interdisciplinarity in the Scholarly Life Cycle* (with Aagje Swinnen).

Participants:

Angelika Braun received her academic training in linguistics and phonetics at Marburg and Cologne Universities. She took her Ph.D. in linguistics and phonetics in 1988. From 1986 to 2000, she worked in forensic phonetics full-time, first at the Forensic Science Laboratory of the Federal Criminal Bureau (BKA) and then as head of the Speaker Identification Section of the Forensic Science Laboratory of the State of North Rhine-Westphalia. In July 2000 she earned her post-doctoral degree (*Habilitation*) from the University of Marburg. Later that year, she became an associate professor at that university. From 2009 until her retirement in 2021, she held the post of Full Professor of General and Applied Phonetics at Trier University. At present, she is still doing research and forensic casework. Her main research interests are in forensically related areas, e.g., hesitations and verbal irony as well as the history of (forensic) phonetics.

Herbert Masthoff held the position of Research Fellow at the Department of Phonetics at the University of Trier from 1990 to 2017. In that capacity, he taught a variety of courses in General, Applied and Acoustic Phonetics. He has supervised Master's theses as well as doctoral dissertations and his own work emphasizes topics relevant to Forensic Phonetics. He presently is a lecturer at the University of Trier and continues doing forensic phonetic casework as an expert witness, which he started more than thirty years ago. His expertise is in forensic phonetic applications of phonetics, specifically in voice analysis, speaker profiling, voice comparison and speech enhancement. He has been a member of the International Association of Forensic Phonetics and Acoustics since 1994.

Maartje Schreuder works at the intersection of (legal) psychology and forensic linguistics, in this way combining her linguistic background (Ph.D. Groningen University, 2006) with her current work as a lecturer at the Faculty of Psychology and Neuroscience at Maastricht University. Her research interests are in forensic speech analysis, earwitnesses, and cognitive biases in forensic experts' work, the last topic strongly relating to her case work as an expert witness for The Maastricht Forensic Institute.

Peter French spent the early part of his career working in traditional dialectology, child language development, language and education and conversation analysis. Since the mid-1980s he has worked in phonetics and acoustics, concentrating on forensic applications. He is chairman of an independent forensic speech and acoustics laboratory (JP French Associates) and has acted as an expert witness in legal cases in jurisdictions across the world. His main research interests are in modelling individual speaker characteristics, forensic speaker comparison (human and automatic) and uses of phonetic and acoustic analysis for determining unclear speech content. For fifteen years, he was President of the International Association for Forensic Phonetics and Acoustics. He is Emeritus Professor of Forensic Speech Science at the University of York, Visiting Professor at York St. John University, and a member of the Centre for Forensic Phonetics and Acoustics Executive Board at the University of Zurich.

05

REFERENCES

REFERENCES

- Bijsterveld, K. (2021). Slicing Sound: Speaker Identification and Sonic Skills at the Stasi, 1966-1989. *Isis*, 112 (2), 215-241.
- Bijsterveld K. & Kvicalova, A. (Eds.) (2023). *Forensic Voices: Cultures of Sonic Identification and Detection in the West*. Special issue of *Sound Studies*, 9(2), forthcoming
- French, P. (2017). A Developmental History of Forensic Speaker Comparison in the UK. *English Phonetics*, 21, 271-286.
- French, P. & Gold, E. (2019). International practices in forensic speaker comparison: Second survey. *International Journal of Speech, Language and the Law*, 26(1), 1-20.
- Gold, E. & French, P. (2011). International practices in forensic speaker comparison. *International Journal of Speech, Language and the Law*, 18(2), 293-307.
- Gold, E. & French, P. (2019). International practices in forensic speaker comparisons: second survey. *International Journal of Speech, Language and the Law*, 26(1), 1-20.
- Kvicalova, A. (2023). Sound on the Quiet: Speaker Identification and Auditory Objectivity in Czechoslovak Fonoscopy, 1975. *Technology and Culture*, 64(2), 379-406.
- Laver, J. (1994). *Principles of Phonetics*. Cambridge: Cambridge University Press.
- Maas H. (2018). The Method of the Witness seminar. *History of Political Economy*, 50(3), 571-577.
- Morrison, G. S., Sahito, F.H., Jardine, G., Djokic, D., Clavet, S., Berghs, S. & Goemans Dorny, C. (2016). INTERPOL survey of the use of speaker identification by law enforcement agencies. *Forensic Science International*, 263(June), 92-100.
- Nicholls E. J. (2020). The witness seminar: A research note. *Qualitative Research*, 00(0), 1-8.
- Potter, R. K, Kopp, G. A. & Green, H. C. (1947). *Visible Speech*. New York: Van Nostrand.
- Tansey E. M. (2006). Witnessing the witnesses: potentials and pitfalls of the witness seminar in the history of twentieth-century medicine. In: Doel R. E. and Söderqvist T. (Eds.) *The Historiography of Contemporary Science, Technology, and Medicine: Writing Recent Science*. Oxon: Routledge, 260-278.

**COMPARING VOICES:
SPEAKER IDENTIFICATION WITNESS SEMINAR**

Transcript of a Witness Seminar on Speaker Identification,
Maastricht University, the Netherlands
(in person, 17 June, and online 9 November 2022).

Anna Kvicalova,
Karin Bijsterveld

Published by Maastricht University

ISBN 978-90-9037363-8