

DESIDERIA – A CASE REPORT ON IDENTIFICATION BY VOICE LINE-UP

Michaël STENBERG

National Laboratory of Forensic Science, Linköping, Sweden

ABSTRACT: The present paper reports on a successful case of identification by voice line-up, where an excerpt from a recorded police interview with a suspect was matched with recordings of semi-spontaneous speech elicited from five foils. Before selecting the foils, interviews were held with the victims in a so-called cognitive style. The “cognitive interview” technique, reputed to make people recall things they seem to have forgotten, is often practised in making “E-Fit” or other composite images to assist in facial identification of perpetrators of crime.

KEY WORDS: Voice line-up; Semi-spontaneous speech; Speech elicitation; “Cognitive interview”.

Problems of Forensic Sciences, vol. XLVII, 2001, 184–189

Received 19 December 2000; accepted 15 September 2001

INTRODUCTION

The arranging of voice line-ups for identifying purposes is a relatively sparsely used forensic tool. Consequently, there is no generally accepted model for the procedure. Several issues are continuously subjected to discussion, e.g., what kind of speech material to use, how to select foils and how to administer the line-up. In Sweden, The National Laboratory of Forensic Science (SKL) in Linköping has taken on the responsibility for conducting such procedures in co-operation with law enforcement agencies throughout the country. The present case is the first one to have been entirely carried out under the auspices of the SKL since the establishment of its Speech Analysis Group in 1995. In Sweden, visual line-ups have been quite widely in use for a long time. However, they suffered a considerable loss of prestige following the revelation to the public of the unprofessional way in which the police investigators had conducted a line-up when dealing with the case of the murder of the Swedish Prime Minister Olof Palme in 1986. As a consequence, guidelines for visual line-ups were promulgated. Therefore, a thorough scrutiny is to be expected whenever line-ups are brought into court, be they auditory or visual ones. The present case was retried in the court of appeal, the defending party having appealed against the judgement of the district court.

BACKGROUND

On the night of the 24th to the 25th of November 1999, a money transport van belonging to one of the major security companies in Sweden left the city of Örebro bound for Västerås, another city some 100 kilometres away. Its crew consisted of two female security guards who had travelled this route several times before. Extensive road repairs were going on long since, and shortly after the departure, the van reached a speed-restricted stretch. All of a sudden, the driver had to jam on the brakes to avoid running into a private car in front of her, which had stopped dead. Hardly had the van come to a standstill when a bang was heard and the door was torn open by a man (henceforth “Robber X”) who told the driver to make room and then took over the driver’s seat and set off. The security guard who had been driving the van (henceforth “Guard A”) discovered that her left arm had been injured by a bullet. The other robber (henceforth “Robber Y”) followed in a car which proved to be the one that had stopped in front of the van. After a while, both vehicles left the motorway for by paths, ending up in a field where the two robbers tried to break open the van, the security guards watching them. However, since the van was of recent design, unfamiliar to the robbers, they did not manage to take possession of the money in it, but had to leave the place empty-handed.

PREPARATION OF THE LINE-UPS

The ensuing preliminary investigation gave rise to suspicions of four men, previously known to the police, two of whom were supposed to be Robbers X and Y. On the 16th of December the Örebro county criminal investigation department applied to the SKL for assistance in arranging voice line-ups, and on the following day a forensic phonetician carried out interviews with the crew of the van. This was done in a so-called cognitive style: instead of showering the interviewees with questions, they were asked to recall the setting of the crime and give their accounts of what happened at their own pace, following their own association paths. Only when their memory seemed totally exhausted, they were requested to expound on certain essential matters. This “cognitive interview” technique, reputed to make people recall things they otherwise would seem to have forgotten, is often practised in making “E-Fit” or other composite images to assist in facial identification of perpetrators of crime.

During the interview, Guard A declared that she probably would be able to recognise Robber X if given the opportunity to hear him speak again. The other guard (henceforth “Guard B”), who had been sitting to the right in the

van, did not make any declaration of this kind. None of them could give an unequivocal description of Robber X's dialect. Tracing on maps of Sweden, they tried to narrow down the portion of the country from which he could possibly derive his origin. The outcome was a rather big area, mainly comprising the western part of south Sweden. Guard A, however, remembered that two days after the assault, she happened to watch "Bingo-Lotto", a popular TV programme. It had then struck her that the compère of the programme was speaking the same dialect as Robber X. This direct comparison was judged by the phonetician to be a more positive indication than the somewhat vague determination first given. Furthermore, the guard described his voice as "neither clear nor dark", adding "absolutely not dark". Judging not only by his voice but also by his appearances, she estimated his age at between 25 and 30 years. In spite of the fact that the statements of Guard B *per se* hardly could justify the arranging of a line-up procedure, it was decided – at the request of one of the police investigators – that line-ups should be organised with both victims. "Voice line-up: A Practical Approach" by A. P. A. Broeders and A. G. van Amelsvoort, in its January 2000 version, was adopted as a set of guidelines.

It was obvious that the dialect of the compère belonged in Gothenburg (Göteborg) or its closest surroundings, which meant a narrowing down of Guard A's (as well as Guard B's) first description. A subsequent check-up with the compère in person proved him to be born and bred in Gothenburg itself. Since one of the suspects was from that city, there was no question about where to hunt for foils. A police commissioner in Gothenburg was contacted and was asked to assemble at least ten young men who had grown up and spent most of their lives in Greater Gothenburg. Shortly afterwards, the phonetician and the policeman who had questioned and recorded the suspect headed for Gothenburg. Using the same tape recorder and microphone, as well as the same type of tape, recordings were made of ten foils, aged between 29 and 42.

Since the sole recorded speech material from the suspect was a police interview, it would have been inappropriate to have the foils speak about subjects at their own option. A fair solution seemed to be asking them to render a passage from the recording of the suspect fluently though not verbatim. A passage of coherent speech, about twenty seconds long, was selected for the purpose. The foils were instructed to acquaint themselves with it by reading a transcription of it, then to practise telling its content in their own wordings. The duration was to be kept at around twenty seconds. No manuscripts whatever were going to be allowed during the recording. Before being recorded, each of the foils had to make several trials before the phonetician and the policeman. Some of them were found to need additional practice.

In order to get the speech situation as authentic as possible, each recording was preceded by the policeman making the same statement as in the interview with the suspect, thus eliciting an answer from the foil.

TESTING THE LINE-UP

One of the suggested foils was eliminated at an early stage, because he turned out to be born and have grown up rather far from Gothenburg and have spent only a short time there. A test tape was compiled, consisting of the recordings of the remaining nine foils, plus a corresponding excerpt from the recorded interview with the suspect. Then five women employees at the SKL acted as test listeners with the purpose of finding out whether any of the recordings stood out from the others in a way that might give cause for disqualifying it. Written instructions were distributed immediately before the beginning of the test. One test listener found that two of the samples sounded somewhat “studied”. In only one of the cases, however, the remark was assessed enough well-founded to justify exclusion. Anyhow, in order to bring down the number of foils to the required five, three more were eliminated, one of them due to misplaced stress.

According to Swedish legal usage, the defence has a right to have a say in matters of this kind. Since the suspect had expressed a wish to be placed “just before the middle of the line-up”, his sample was assigned number three in both line-ups. This done, the foils were randomly distributed, differently in the two line-ups.

OPERATING THE LINE-UPS

On the 2nd of February 2000, the victims were summoned to the Police Headquarters at Örebro to take part in the line-up procedure. Each of the line-ups was administered by a detective inspector that had never met the suspect. Written instructions were distributed to the inspectors, who were also orally briefed. Before setting to work, they had to listen to the line-ups individually in order to make certain that they did not happen to recognise any of the speakers.

Prior to the line-up presentation, the victims received instructions in writing and were given the opportunity to ask questions. For practical reasons, the samples had to be played back from compact cassettes. In order not to let the victims know in advance how many samples they were to hear, the cassettes were hidden and taken out one by one. To familiarise the victims with the procedure, a recording of a test speaker was played immediately be-

fore the first sample of the line-up. This recording had approximately the same length and content as the samples in the line-up, but was presented in Malmö dialect, clearly distinct from that of Gothenburg.

RESULTS

Guard B was the first one to be presented with a line-up. She listened to all the samples without any comment or visible reaction. After completion of the playback, she declared that she thought the voice in sample no. 2 “was awfully alike the voice of the robber”. Later she said that no. 2 was the most similar one. When the procedure proper was over, she was asked whether the dialect in the line-up was that of Robber X. Her answer was yes. Before the court Guard B stated that she could not recognise any voice.

Guard A was then presented with a line-up in which the suspect was still no. 3, but where the foils were placed in a different order. During the playback of no. 3, she uttered: “It sounded rather alike... but it gets so dark at the end of the recording”. After the playback of no. 6 (the last sample) she said: “He didn’t say the same thing... for it was just that thing with the headlamp I would have liked to hear”. She asked to hear no. 3 once again and was allowed to do so. After the playback she stated, “I think it is no. 3”. A short while later she declared: “Well, I believe in no. 3”. Before the court Guard A said as follows (after the prosecutor having asked her what it was like to listen to the recordings): “It was difficult. When hearing person no. 3 my pulse speeded up and I knew it was him. You felt that it was that voice... it was the overall impression, but he has a voice that is somewhat clearer. I’m 99 per cent certain that it was that voice that was at my side”.

The district court found that through the voice line-up and through the statements given by Guard A, there was good support for the presumption that the suspect was identical with Robber X, the one who shot her in her arm.

Robber X was sentenced to 10 years’ and Robber Y to 9.5 years’ imprisonment. The sentences included other crimes like theft, receiving stolen goods and violation of the arms regulations. The defence appealed against both sentences.

In the beginning of June proceedings took place in the court of appeal. One of the two defence lawyers, who in the district court had had no questions to ask regarding the line-up, now pretended that before ever having heard the suspect speak, he had made a correct guess at which sample originated from him. This lawyer thus insinuated that there might be a lack of naturalness in the foils’ way of speaking. In its sentence, the court of appeal did not attach great importance to the line-up. The plea grounds for this were two: Firstly, the court members found the speech sample from the sus-

pect clearly deviant from those of the foils. Secondly, according to information received, the victims had been informed – by the police – about the outcome of the line-up procedure before they appeared in court.

DISCUSSION

The age of the foils – those actually used were between 34 and 42 – may seem a bit high, considering the suspect's 28. However, in Broeders' and van Amelsvoort's guidelines, the main stress is laid on perceived age. Since none of the test listeners found any reason for objections in this respect, it could be presumed that the foils selected were all right for the purpose. Still, time permitting, an even better way of testing the line-up would have been having a panel of at least 30 persons listen to the samples and guess which one represented the suspect. Any agglomerations of guesses beyond chance would then serve as a memento that a revision of the line-up is called for.

The borderline between what sounds like spontaneous and non-spontaneous speech is of course very difficult to draw in a case like this. Even if it can be supposed that the judgement of the court of appeal might have been influenced by the presumption of the lawyer, which could very well have been a downright lie, there seems to be reason for questioning the use of anything but purely spontaneous speech in line-up contexts. The district court's ready acceptance of the elicited speech as spontaneous-like might as well be biased by an over-confidence in experts. Two options seem to remain for the serious forensic phonetician: either perform an elaborate testing procedure with a big panel of test-listeners prior to the line-up procedure, or – maybe even better – refrain from using any kind of elicited speech.

Given the fact that not only voice characteristics, but also linguistic features like dialect or accent in all its aspects (pronunciation, vocabulary, grammar), speech rate, phrasing etc. play an important part in the identification process, it seems that speech line-up would be a more adequate designation than voice line-up.

References:

1. Broeders A. P. A., Rietveld A. C. M., Speaker identification by earwitnesses, [in:] *Studies in forensic phonetics*, Braun A., Köster J.-P. [eds.], WVT Wissenschaftlicher Verlag Trier, Trier 1995.
2. Hollien H., *The acoustics of crime*, Plenum Press, New York, 1990.
3. McGehee F., The reliability of the identification of the human voice, *The Journal of General Psychology* 1937, vol. 17, pp. 249–271.
4. McGehee F., An experimental study in voice recognition, *The Journal of General Psychology* 1944, vol. 31, pp. 53–65.