Neutral Citation No. (2002) NICA 20

Ref:	NICB3173
Delivered:	19.04.2002

IN HER MAJESTY'S COURT OF APPEAL IN NORTHERN IRELAND

THE QUEEN

v

ANTHONY O'DOHERTY

NICHOLSON LJ

Introduction

On 15 October 1997 the appellant was convicted on indictment before His Honour Judge Smyth QC and a jury of two offences:-

(1) Aggravated Burglary contrary to section 10(1) of the Theft Act (Northern Ireland) 1969;

(2) Causing Grievous Bodily Harm with intent, contrary to section 18 of the Offences Against the Person Act 1861. These offences were alleged to have been committed on 5 December 1995. He was sentenced to 12 years' imprisonment.

He applied for leave to appeal against conviction on both counts of the indictment.

On 30 October 1998 the Court of Appeal refused leave to appeal.

In the course of their judgment they said:-

> "The learned trial judge in directing the jury described the evidence linking the applicant with the offences as containing four planks.
>
> 1. The evidence of Detective Sergeant Johnston of his recognition of the voice of the male caller to ambulance control as that of the applicant.

2. The use of a mobile telephone stolen by the burglars to call the applicant's house three times within a period of thirty minutes after 2 am on the night of the burglary (which occurred about 10 pm).

3. The evidence of Mrs McClelland [the expert witness called by the prosecution on voice identification] to the effect that it was highly probable that the applicant was the male caller to ambulance control.

4. The comparisons that the jury themselves could be invited to make from having heard the tape of the call to ambulance control and the voice of the applicant as he gave evidence in court".

The Criminal Cases Review Commission referred the conviction to the Court of Appeal by its decision of 27 September 2001 under Section 10 of the Criminal Appeal Act 1995. An application for the reception of fresh evidence was made on behalf of the appellant on 13 December 2001. We granted the application and gave bail to the appellant on 20 December 2001. An application for the reception of further fresh evidence was made in February 2002 and was also granted. The reference was heard between 25 and 27 February 2002.

The Fresh Evidence

We heard fresh evidence in relation to the first, third and fourth planks of the evidence linking the appellant with the offences. Evidence was given by Dr Nolan on behalf of the appellant and by Dr French and Mrs McClelland on behalf of the prosecution. Reports of Dr Kunzel and Professor Bull were agreed to be read into the record on behalf of the appellant.

Dr Nolan is Reader in Phonetics at the University of Cambridge. His reports of 24 April 2000, 18 March 2001, 11 July 2001 and 19 February 2002 were agreed to be read into the record. In his summary of his findings on 24 April 2000 he stated:-

"As far as the identifications made by the police officers are concerned, I believe these are highly problematic:

Auditory identification of speakers known to untrained listeners, contrary to popular belief, yields high error rates even under ideal listening conditions.

While certain speakers may be distinctive, and accurately identified, other speakers may be particularly prone to confusion, with the danger of 'judicially fatal' errors.

There is a very serious danger of inadvertent psychological bias when speaker identifications are carried out by police officers involved in investigating a crime.

Auditory judgments (whether by phoneticians or laymen) to the effect that the 999 caller [to ambulance control] sounds like Mr O'Doherty have no weight against demonstrable acoustic differences for which the only justifiable explanation is that the 999 call is made by a different speaker.

My own analysis shows that there are clear and quite obvious differences in acoustic pattern and even in auditory pattern, between Mr O'Doherty's speech and that of the 999 caller. I cannot understand why these have been overlooked or set aside other than the fact that Mrs McClelland [who gave evidence in court for the prosecution] did not, apparently, carry out a thorough instrumental acoustic examination of the material. There is no justification for the conclusion that Mr O'Doherty is 'very likely' to be the 999 caller. The acoustic evidence shows, on the contrary, that Mr O'Doherty's voice is incompatible with that heard on the 999 caller.

The jury was neither in a position to apply the techniques required to reach a reliable view on speaker identity, nor would they have been free of inadvertent psychological bias.

It must be said at the outset that it is rarely if ever possible to achieve certainty in identification by voice. A person's voice is quite unlike a fingerprint. A fingerprint is unchanging and unique, a voice is variable, and it has not been scientifically proven how extensively features of the voice are shared among members of a population. Nonetheless a phonetician, using trained auditory skills and acoustic analysis can often make observations which can weight for or against the possibility of two recordings being from the same individual.

I have consistently argued (eg Journal of Linguistics, 1991 and J. Gibbons ed. Language and the Law (1993)) that while auditory phonetic analysis is good at telling us whether two samples have the same accent, once it is established that two samples have the same accent, and generally similar voice quality, only

> quantitative acoustic analysis can go further and come anywhere near to determining whether the two samples of the same accent come from the same individual. One reason for this is that our hearing system has to set aside a lot of acoustic differences between individuals (eg people with different shapes and sizes of mouths) in order to hear what is being said, and these differences are among the features which can show up on an objective instrument – based acoustic analysis. The position that quantitative acoustic analysis is required before any conclusions can be reached stronger than 'speakers with the same accent with rather similar voice quality' is fully accepted in those countries (notably Germany and the Netherlands) where forensic phonetics have been integrated into the Government forensic science service.
>
> The telephone call from the Ambulance Service provided a sample of 16 seconds of speech from the male speaker, most of which was intelligible but the acoustic signal was significantly degraded. In the 999 caller's speech the information above 2K hz is already very weak. When one considers that the information which we make use of in perceiving speech extends at least up to 8K hz then it is clear that the sample of the male 999 caller is seriously degraded. Most of the acoustic information about fricatives (hissy sounds like 'f', 's', 'sh' and 'th') and some of that about plosives (sounds like 't' and 'k') occur above 3K hz and certainly above 2K hz. It is disingenuous of Mrs McClelland to claim, as she did in her evidence at the trial that 95% of the consonants are available for analysis in the sample. None of the 's', 'f', 'th', 'sh', 'z', 'v', 'ch', 'j' consonants and a few others will be present in a way which would allow them to be heard as such … let alone subjected to a detailed comparison for 'between speaker' differences.
>
> There is a broad similarity in terms of accent between the 999 caller and Mr O'Doherty … since the acoustic information that could signal differences in some sounds is missing from the 999 tape it would be incautious to claim that they have the same accent. And of course many people share an accent.
>
> … A striking difference is to be heard in the pronunciation of the word 'person'. It is a salient auditory difference and must weight against the conclusion that the speakers are the same. Mrs McClelland either did not pick it up, or if she did notice it, did not see the need to explain why she set it aside".

Dr Nolan discussed other aspects of Mrs McClelland's evidence and then proceeded to set out his acoustic analysis and findings by way of graphics. He expressed surprise that the

jury were allowed to listen to the tapes and, in effect, make up their own minds. The members of the jury were, he said in his report, naturally not in a position to carry out instrumental acoustic analysis, nor even to bring to bear the kinds of phonetic and linguistic analysis which the experts used. The jury members were already inevitably under psychological bias from the very fact that Mr O'Doherty was in the dock and had been confidently identified by a police officer of his acquaintance. In this frame of mind it was not surprising if they heard a faint, noisy, partially unintelligible telephone recording of a broadly similar voice as being that of Mr O'Doherty.

In his evidence to the Court he agreed with the evidence of Dr Baldwin, called on behalf of the prosecution in R v Robb (1991) 93 Cr App R 161 referred to in the judgment of the Court of Appeal at p 165 that:

> "the great weight of informed opinion, including the world leaders in the field, was to the effect that auditory techniques unless supplemented and verified by acoustic analysis were an unreliable basis of speaker identification".

Auditory phonetic analysis tells one principally about the dialect or accent of the speaker; quantitative acoustic analysis enables one to examine the differences in the acoustic properties of the speech which depend on the individual's vocal tract, mouth and throat, he said.

He was not aware of any country in continental Europe where expert opinions on voice identification were based exclusively on auditory techniques and the standard method involved a combination of both methods. The speech signal in the 999 call had lost a lot of its high frequency energy which normally does happen over the telephone but in his experience the loss in this case was rather greater than the average telephone call. It meant that a lot of the sounds that could potentially be different between two speakers were simply not there … voice quality was changed by the telephone and it was much more difficult to recognise a voice over the telephone than in a good quality recording or in the presence of a person. Roughly half the

sounds of English are consonants and probably half of these are the ones seriously affected by the telephone.

He explained in some details the spectral analysis of the speech and the spectrograms which were to be found at the end of his report of 24 April 2000. Each mark (which looks almost like a distorted bar code), changing slightly up and down, indicated a resonance of the voice which is sometimes called a formant. He analysed the word "person" in the 999 call and the word "person" as spoken by Mr O'Doherty in a taped police interview. (The speaking voice ranges from about 100 hz or 0.1 of a K hz up to above 8K hz). The second and third resonances were primarily dependent on the shape and size of the mouth and throat and in some sounds the nasal cavity as well. The pitch of the voice may go up in a stressful situation [a 999 call] but he did not think the voice tract of the mouth and throat would change in shape or size. There is an element of variability inherent in all speech production but in the case of the word "person" the variation was greater than one could reasonably expect from random variation within the acoustic output of a single speaker. Ideally one would look at many, many examples of the same word but there were very few examples of any word on the 999 call. The acoustic analysis corresponded to what he would expect for one speaker who pronounced the "r" in "person" and one who didn't. Mr O'Doherty did pronounce it. The speaker on the 999 call did not. The combination of differences in the formants pointed to two speakers. The differences were of a magnitude which, he thought, greatly exceeded anything one would expect of a single speaker.

He examined the words "an ambulance". The primary conclusion to be drawn was that the speakers of these two words on the 999 call and in the taped police interview were different speakers. The recurring pattern in all the comparisons was that Mr O'Doherty had a very low first resonance whereas the 999 caller did not.

The 999 call contained, apart from "aha", only one example of any word and so any form of speaker identification had to proceed very very cautiously.

The difference in pattern in “person” was a separate observation from the difference in “an ambulance”. There was a difference of pattern in “Hello” and “Aha” which suggested to him that the most likely explanation was that these words were produced by different speakers. He also considered that the name “Moore” as spoken on the 999 call was different from “Moore” as spoken by Mr O’Doherty.

There were three occurrences of “Aha” in the 999 call and Dr Nolan randomly picked out three occurrences of “Aha” in the police interview tapes with Mr O’Doherty. Each time Mr O’Doherty had a very low first formant and a relatively high second formant whereas the 999 caller had a higher first formant and a lower second formant.

Dr Nolan found an accumulation of differences, the best explanation for which and probably the only explanation for which was, he said, that they were produced by different speakers.

As to police identification of the voice, he said that voice identification was very difficult and so any slight bias in the mind of the listener can have an effect.

To ask the jury to make up their minds suggested strongly that they should actually disregard the expert evidence. Simply because the experts disagreed does not mean the experts’ evidence was of no value. The true value of that evidence was that identification is very difficult to achieve and the experts at the trial came to different conclusions. Secondly, when the jury has been sitting in the presence of somebody accused of a crime it must be almost impossible for them to approach the task of identification in an unbiased way. Thirdly, there is no guarantee that the jury members were appropriately familiar with the kind of voices involved in this case. Fourthly, the whole point of identification line-ups was to provide some statistical protection to an innocent suspect who looks or sounds a little bit like the perpetrator of the crime. The jury were in effect given a line-up of one and asked “Is this the person or not”. Sixthly, the jury were not given acoustic analysis which gives one complementary information.

His view was that a jury should not be asked to listen to the tape or allowed to use it to make any kind of identification.

In cross-examination he stated that a phonetician would be quite a bit more discriminatory then a lay person. If accent features, which are extremely rare in the population, are identified they may be able to come to a conclusion as to the likelihood whether or not the suspect sample is attributable to the accused.

It was unlikely that on the basis of auditory analysis alone, experts could come safely to a conclusion that "it is highly likely that these two samples are from the same person". They could give the opinion that the accent is the same and the voices sound generally similar. It would be a safe position for auditory phoneticians to take that there has to be some rare individual feature before an opinion can be given as to the likelihood of the suspect sample being attributed to the accused. Phonetic analysis does not purport to be a tool for describing the difference between one speaker and another, the differences which arise from the vocal mechanisms. The way in which we hear will fail to distinguish quite a number of the features which are important in deciding whether samples came from two speakers or one.

The auditory phonetician can very helpfully say: It is fully possible that these two samples came from the same speaker. To go beyond that one needs to find an absence of acoustic differences. Going from "possible" to "more likely than not" is going from the realm of possibility to probability and that is where, said Dr Nolan, the phonetician should normally draw the line. The scales used by Mrs McClelland and Dr French confused possibility and probability in a way which was not supportable. He was speaking personally but was the chairman of the committee which took on the task of investigating these scales. On the scale of – 5 to 5 a purely auditory phonetician could reach "zero" but he would be happy for the phonetician to say "fully possible" in order to indicate that every auditory avenue had been explored.

Interpretation can give rise to different conclusions but "subjective" is not an appropriate word to use. He agreed to the phrase "a degree of subjectivity". He said that he was very cautious about the word "probability" unless there were idiosyncrancies. If two samples have a very very low pitch and a dozen or so features which are very much towards the edge of the population then one can begin to talk about probability and say that it is unlikely to find two samples that match in those relatively rare values except where they are both spoken by the same person.

Auditory analysis gives you a description of the accent and in addition you can make observations about voice quality, possibly rhythm and certainly intonation. The judgments on those aspects of the voice are made in an impressionistic way, much as the lay person would make them, but there is available a systematic analysis of voice quality. He had never seen it referred to in a report on voice identification. He acoustically analysed those words which were clearest on the 999 call, bearing in mind that it was very poor material. To seek to explain an acoustic difference on the basis that Mr O'Doherty was more nasalised in interview than he would normally have been was an attempt to explain a major acoustic difference after the event. He considered that "hello" was a valid candidate for analysis in speaker identification. He understood Dr French to have completely agreed with his analysis and merely to have found a different way of interpreting the substantial acoustic difference.

In answer to the Court he said that perhaps 5 or 10 per cent of voice experts stick to the view that nothing can be added by acoustic analysis. The practical difficulties in using acoustic analysis were much greater when the case of Robb was decided than they are today … In 1989, 1990 one needed very expensive equipment. The equipment is better now. The time had come to move on where the matter is one involving the identification of a single individual.

He thought that he was in a much better position by the end of his listening to the tapes to make an identification than the police officer who had spoken ten or twelve times in a year to

the appellant. He was listening essentially only for the aspects which were to do with the identity of the speaker and not at the same time holding a conversation as presumably the police officer was. There was a serious danger of bias in the police officer's mind if his only contact with Mr O'Doherty was in the context of crime.

He repeated his belief that the jury would have been in a very bad position to make an assessment as to identification. During the course of the trial they would principally have been attending to the content of what Mr O'Doherty said. They may be unaware of the difficulties which the telephone imposes on voice identification because of the loss of acoustic information. Their bias in favour of finding the defendant to be the voice on the tape was by far the strongest. They weren't able to hear again and again. They had no training as to voice quality, pitch and intonation. The telephone recording was bad. If what they were doing was trying to remember the voice of Mr O'Doherty which they heard under the stressful and distracting context of the court case and later listen to the 999 tape, that would be disgraceful, he thought.

It was pointed out to him that the jury asked at the end of the case for the tape to be played again and it was played again. He said that if it was a courtroom such as the one in which he was giving evidence they should at least have had the benefit of hearing the tape directly over headphones. It seemed negligent not to have allowed them to hear adjacently some sample of the defendant's voice to act as a comparison. One of the factors that makes one more error prone is the time elapsed between hearing the voice and having to make an identification. The jury should not be in the position of having an identity parade of one. But if it is going to happen then all precautions should have been taken to make the listening conditions as good as possible and to warn them of the very special difficulties involved in making such identification, including the dangers of the telephone speech. The warning should have several factors in it, not just a warning about the possibility of error in lay identification but probably a warning about

the poor quality of the material and the fact that this may lead to difficulty in deciding who the speaker is.

This concluded Dr Nolan's evidence.

The Crown then sought leave to call Dr French and Mrs McClelland and this was granted.

Dr French's report was agreed to be read into the record. He has a distinguished academic background and experience in giving forensic evidence in court more than 200 times throughout the world.

So far as the 999 call was concerned, he said, one would expect a telephone call in that context to reflect a higher fundamental frequency value than in police interviews. The disparity found was what one might expect. He described what an auditory phonetician would be looking for when carrying out his or her examination. They would be particularly interested in any features of an unusual and personally distinguishing nature. His advice to the Crown would be that if they only had auditory and acoustic analysis of a telephone call as evidence of identification against a defendant, the prosecution should not proceed.

As to whether or not Mr O'Doherty's voice was on the 999 telephone call expressed on a scale of 0 to plus 5 and 0 to – 5, he would put it at the first positive point, namely, "rather more likely than not". "Probable" for him would be two points further up the scale. He found similarities in voice quality, nasalisation and breathiness, duration of tonic syllables carrying the main sentence stress, duration of words, main syllables not having pitch movement on them. He felt that the voice quality, the rhythmic and intonational features made a relatively distinctive voice in the interviews and in the 999 call. He factored in the dissimilarities. He discussed his findings as compared with Dr Nolan's findings. The differences were not of a magnitude that would cause him to eliminate Mr O'Doherty as the speaker on the 999 call, especially when set against the backcloth of similarities. He did not regard the differences as being as significant as

Dr Nolan did. He would normally discard the word "hello" but accepted that it was different in the 999 call and in the police interview when the appellant read aloud. The first formant was significantly different but for reasons which he gave, he discounted this. He agreed with Dr Nolan about "an ambulance" but said that some people with low nasalised voices in interview tend to speak in a slightly less nasalised way in other less subdued circumstances.

He agreed largely with Dr Nolan about the "ahas" but said that the differences have to be set against the background of similarities found. He discussed telephone distortion in a similar way to Dr Nolan.

He thought there were dangers in playing tapes to juries and had expressed this view in print after an Appeal Court ruling that the jury should hear the tapes. (Presumably R v Bentum referred to later in this judgment).

He routinely carried out acoustic analysis. Auditory analysis and acoustic analysis provided cross-checks against one another. That would be best practice. It would be the general view.

He only knew one expert left in the UK who did not do acoustic testing. In the light of the results that he was able to obtain there was a significant modification made by him to the conclusion which Mrs McClelland had reached. He did not adopt the approach of Mrs McClelland that "comparison is not carried out on the basis that the reference is being compared with an open population of voices".

The difference between him and Dr Nolan was one of interpretation rather than fact. There were no population statistics against which auditory or acoustic analysis can be tested. He was firmly of the view that voice identification should not be undertaken by a layman.

Mrs McClelland referred to her report of 19 March 1987 and her evidence at the trial. She also had written responses to Dr Nolan's reports which were agreed to be read into the record. She had examined data in approximately 800 cases. She did not carry out a full acoustic

analysis of the tape recordings for the purposes of giving evidence at the trial. She did not consider formant analysis as contributing in any principled way to decisions on identity or non-identity between speakers.

Since the trial she re-examined the tapes. She did not measure the formants. She remained of the same view as she did at the trial that it was highly probable that the voice in the 999 telephone call was that of Mr O'Doherty.

In cross-examination she said that formant measurements were unreliable and unsafe. She made a statement of evidence for the trial in which she said that there were thirty seconds of recording on the 999 call which were available for analysis; she revised this to 20 seconds at the trial and now accepted that there were 16 seconds. She agreed that when she told the jury that there were 95 per cent of consonants and 90 per cent of vowels available for expert analysis, she unintentionally misled the jury. In phonetic terms a lot less information was available. In 1996 she was carrying out very little acoustic analysis. Since 1997 she carried out acoustic analysis in all cases except when there was a closed set of suspects and she was being asked to say which voice was which. She did not carry out a lot of formant averaging and measurement. She agreed that it was the generally accepted practice within the United Kingdom and beyond to do so. At the time of the trial she told the jury that the amount of data in the 999 telephone call was such that acoustic analysis would not be very useful. In her report commenting on Dr Nolan's report she said that there were not any population statistics against which the measurements could be tested. She did not say that it was her practice to carry out acoustic tests because she had ongoing uncertainty as to their value, as to what they actually mean. She accepted that there were no population statistics to validate any conclusion reached as a result of auditory analysis. In the latter field one is on very subjective ground. It was a surprise to her to read Professor Kunzel's report and see that the scientific and legal communities throughout continental Europe insisted on acoustic as well as auditory tests. She was a sceptic about

acoustic analysis. She was surprised that Dr French dissociated himself from her approach that the comparison was not carried out on the basis that the reference sample voice was being compared with an open population of voices.

She had become increasingly sceptical on the topic of formant measuring. This completed the fresh evidence which we received.

We are very grateful to counsel for their very helpful written and oral arguments.

The Principles which apply to the reception of fresh evidence

These have recently between reviewed in R v Pendleton [2002] 1 All ER 525.

In the course of his opinion Lord Bingham said at pp 534, 535:-

> "In hearing any appeal against conviction the Court of Appeal will ordinarily have a considerable body of material before it: grounds of appeal; transcripts of the judge's summing up to the jury and any relevant passages in the evidence and of any material rulings given before or in the course of the trial; plans, photographs and so on. And although the court does not have the jury's reasons, it does have the jury's verdict. From this, some inferences may always be drawn. If the issue is consent, the jury must, to convict, have been sure that the victim did not consent. If the issue is pure identification, the jury must, to convict, have been sure that the evidence identifying the defendant was accurate and reliable. If a proper judicial direction has been given, it will ordinarily be safe for the Court of Appeal to infer that the factual ingredients essential to prove guilt have been established against the defendant to the satisfaction of the jury. But the Court of Appeal can rarely know, save perhaps from questions asked by the jury after retirement at what points the jury have felt difficulty. The jury process of reasoning will not be revealed and, if a number of witnesses give evidence bearing on a single question, the Court of Appeal will never know which of those witnesses the jury accepted and which, if any, they doubted or rejected.
>
> My Lords, Mr Mansfield is right to emphasise the central role of the jury in a trial on indictment. This is an important and greatly-prized feature of our constitution. Trial by jury does not mean trial by jury in the first instance and trial by judges of the Court of Appeal in the second. The Court of Appeal is entrusted with a power of review to guard against the possibility of injustice but it is a power to be exercised with caution, mindful that the Court of Appeal is not privy to the jury's deliberations

and must not intrude into territory which properly belongs to the jury.

Where the Court of Appeal has heard oral evidence under s 23(1)(c) of the 1968 Act (whether pursuant to its own decision, or by agreement, or de bene esse), the evidence will almost always have appeared on paper to be capable of belief and to afford a possible ground for allowing the appeal. By the time the court comes to decide whether the appeal should be allowed or dismissed, it will have heard the evidence, including cross-examination, and any submissions made on its effect. It may then conclude, without doubt, that the evidence cannot be accepted or cannot afford a ground for allowing the appeal. Such was the case, for example in *R v Jones* 33 BMLR 80 at 88, where the court, having decided to receive and having heard opinion evidence from an expert, found conclusive objections to the acceptability of that opinion. The court may, on the other hand, judge the fresh evidence to be clearly conclusive in favour of allowing the appeal. Such might be the case, for example, if a witness who could not be in any way impeached testified, on oath and after all appropriate warnings, that he alone had committed the crime for which the appellant had been convicted. The more difficult cases are of course those which fall between these extreme ends of the spectrum.

It is undesirable that exercise of the important judgment entrusted to the Court of Appeal by s 2(1) of the 1968 Act should be constrained by words not to be found in the statute and that adherence to a particular thought process should be required by judicial decision. Thus the House in *Stafford v DPP* were right to reject the submission of counsel that the Court of Appeal had asked the wrong question by taking as the test the effect of the fresh evidence on their minds and not the effect that that evidence would have had on the mind of the jury (see [1974 AC 878 at 880). It would, as the House pointed out, be anomalous for the court to say that the evidence raised no doubt whatever in their minds but might have raised a reasonable doubt in the minds of the jury. I am not persuaded that the House laid down any incorrect principle in *Stafford v DPP*, so long as the Court of Appeal bears very clearly in mind that the question for its consideration is whether the conviction is safe and not whether the accused is guilty. But the test advocated by counsel in *Stafford v DPP* and by Mr Mansfield in this appeal does have a dual virtue to which the speeches I have quoted perhaps gave somewhat inadequate recognition. First, it reminds the Court of Appeal that it is not and should never become the primary decision-maker. Secondly, it reminds the Court of Appeal that it has an imperfect and incomplete understanding of the full processes which led the jury to convict.

> The Court of Appeal can makes its assessment of the fresh evidence it has heard, but save in a clear case it is at a disadvantage in seeking to relate that evidence to the rest of the evidence which the jury heard. For these reasons it will usually be wise for the Court of Appeal, in a case of any difficulty, to test their own provisional view by asking whether the evidence, if given at the trial might reasonably have affected the decision of the trial jury to convict. If it might, the conviction must be thought to be unsafe.
>
> In some of the authorities, the decision to allow an appeal is closely associated with the decision to order a retrial. This is understandable but wrong. If the court thinks a conviction unsafe, its clear statutory duty is to allow the appeal, whether or not there can be a retrial. A conviction cannot be thought unsafe if a retrial can be ordered but safe if it cannot. It is only when an appeal has been or is to be allowed because a conviction is thought to be unsafe that any question of a retrial can properly arise.
>
> In the present case, as adherence to precedent required, the Court of Appeal formulated a test based on *Stafford v DPP* and other cases in which *Stafford v DPP* had been cited and applied. No criticism of its formulation is made if *Stafford v DPP* itself was correct. Since the principle laid down in *Stafford v DPP* was, in the opinion of the House, correct, the attack made on the Court of Appeal's self-direction in the present case must fail. The foregoing paragraphs it is hoped, make clear the approach which the Court of Appeal should follow".

The application of the principles to this case

A. The evidence of Dr Nolan

Dr Nolan gave evidence that the voice on the tape from ambulance control was not the voice of the appellant. Unless we can discount this evidence completely, we are bound to conclude that the convictions are unsafe. The evidence of Dr French and Mrs McClelland, when weighed against the evidence of Dr Nolan, does not enable us to discount his evidence. In our view there is a reasonable possibility that the voice on the tape was not the voice of the appellant. We do not need to go so far and do not go so far as to say that the voice on the tape is probably not the voice of the appellant. But we consider that the convictions are unsafe for that reason.

One of the planks on which the case against the appellant rested was the evidence of Mrs McClelland to the effect that it was highly probably that the appellant spoke on the tape from ambulance control. We are satisfied that if the jury had heard the evidence of Dr Nolan, borne out by Dr French who was called on behalf of the Crown, that plank of the case would almost certainly have been eliminated.

The decision in R v Robb (1991) 93 Cr App R 161 was to the effect that a witness who was a phonetician and carried out auditory analysis only was well qualified by academic training and practical experience to express an opinion on voice identification. Although the technique relied on by Dr Baldwin, the expert who gave evidence in that case represented a minority view in his profession, he had reasons for preferring to use that technique and he had not, on the facts, been shown to be wrong; the appellant was not in any way unfairly prejudicial by the admission of that admissible evidence the Court of Appeal held. Mrs McClelland who gave evidence for the Crown at this trial gave her evidence based on the same technique as the expert in Robb. We do not say that her evidence was inadmissible, as presented to the Court in 1997 but it is subject to our later comments.

Time has moved on. According to the report of Dr Kunzel, a distinguished academic and forensic expert in voice identification, prosecutors in the rest of Europe invariably present auditory analysis (in which Mrs McClelland specialises) and quantitative acoustic analysis, including formant analysis.

It transpired during the course of the hearing that Mrs McClelland has been carrying out such acoustic analysis since 1997 despite her scepticism and Dr French knows of only one expert in the United Kingdom who relies solely on auditory analysis. Acoustic analysis in 1991 involved the use of expensive computers. Now it can be carried out by the use of an ordinary computer with suitable and readily available software fitted to it".

Despite Mrs McClelland's expression of scepticism we are satisfied, having heard Dr Nolan and Dr French and read the report of Dr Kunzel, that in the present state of scientific knowledge no prosecution should be brought in Northern Ireland in which one of the planks is voice identification given by an expert which is solely confined to auditory analysis. There should also be expert evidence of acoustic analysis such as is used by Dr Nolan, Dr French and all but a small percentage of experts in the United Kingdom and by all experts in the rest of Europe, which includes formant analysis.

We make three exceptions to this general statement. Where the voices of a known group are being listened to and the issue is, "which voice has spoken which words" or where there are rare characteristics which render a speaker identifiable – but this may beg the question - or the issue relates to the accent or dialect of the speaker (see R v Mullan [1983] NIJB 12) acoustic analysis is not necessary. We do not gain any assistance from the decision in R v Gilfoyle (unreported: Court of Appeal, 20 December 2000) to which counsel for the appellant drew our attention.

A second plank (out of four) was "the comparisons that the jury themselves could be invited to make having heard the ambulance control tape and the voice of the appellant as he gave evidence in Court".

In R v Bentum (1989) 153 JP 538 the Court of Appeal in England held that the jury should be allowed to hear any tape recordings for themselves, so that they may form their own judgment of the opinions on voice identification expressed by experts or others claiming to have recognised the voice.

In the present case the Court of Appeal held that the issue was whether the jury considered that the voice on the ambulance control tape was the voice of the appellant. Expert evidence is received when the subject is one upon which competency to form an opinion can only be acquired by a course of special study or experience and it has been accepted that expert

evidence is receivable in cases of voice identification. Expert evidence is rarely, if ever, admitted in cases of visual identification. The tribunal of fact is considered to be in as good a position to assess CCTV footage or video tapes or photographs as witnesses: See R v Murphy and Maguire [1990] NI 306. It seems to us that if evidence of voice recognition is relied on by the prosecution, the jury should be allowed to listen to a tape-recording on which the recognition is based, assuming that the jury have heard the accused giving evidence. It also seems to us that the jury may listen to a tape-recording of the voice of the suspect in order to assist them in evaluating expert evidence and in making up their own minds as to whether the voice on the tapes is the voice of the defendant.

We propose to cite from the judgment of Kelly LJ in R v Murphy and Maguire [1990] NI 306 between p 320F and 327F:-

> "Mr Finegan did not challenge the relevance or authenticity of the heli-tele film or that it was admissible. What was not admissible evidence, he contended, was what the trial judge himself recognised and perceived from his own viewing of the film; this conclusion as to what he himself saw in the heli-teli film was neither admissible evidence nor confirmation of the evidence of the identifying police officers. It was not real evidence. The judge became an unsworn witness and he could not be cross-examined as to what he had observed.
>
> Mr Finegan's conception of real evidence does not however accord with authority. Cross on Evidence, 6th edition p 40 defines the nature of real evidence. There he states under the heading of 'Things or real evidence':
>
> > 'Things are an independent species of evidence as their production calls upon the court to reach conclusions on the basis of its own perception and not on that of witnesses directly or indirectly reported to it ...
> >
> > Although it was devised by *Bentham* and adopted by *Best*, 'Real evidence' is not a term which had received the blessing of common judicial usage. There is general agreement that it covers the production of material objects for inspection by

> the judge or jury in court, but it is debatable how much further the term should be extended'.

Cross goes on to give as examples of real evidence, the court's own observation of the condition of material objects, its own viewing of the physical characteristics of persons, the examination of injuries or their effects in a claim for damages for personal injuries and also the demeanour of witnesses. Then at p 43 of the text under the heading 'Automatic recordings', he states:

> 'Most discussion has hitherto centred on the admissibility of tape-recordings, but this has now been supplemented by a thin trickle of authority on the admissibility of other media such as film, video-tape and computer output. In all of these cases the evidence is real evidence when it is tendered to show what it was that was recorded'.

Murphy in 'A Practical Approach to Evidence' (3rd Ed) also defines 'Real evidence' at p 7:

> 'A term employed to denote any material from which the court may draw conclusions or inferences by using its own senses. The genus includes material objects produced to the court for its inspection, the presentation of the physical characteristics of any person or animal, the demeanour of witnesses (which may or may not be offered or presented to the court by design), views of the locus in quo or of any object incapable of being brought to court without undue difficulty and such items as tapes, films and photographs, the physical appearance of which may be significant over and above the sum total of their contents as such ... What is of importance in each case is the visual, aural or other sensory impression which the evidence, by its own characteristics produces on the court, and on which the court may act to find the truth or probability of any fact which seems to follow from it'.

Later at p 513 (par 16.6.1) the author again in the context of 'real evidence' writes:

> 'The court may look at and draw any proper conclusions from its visual observation of any relevant material object produced before it ... The

> tribunal of fact is entitled to act on the results of its own perception, even where it conflicts with other evidence given about the object ...'.

and at p 515 (par 16.6.5):

> 'The court must, before admitting recordings as evidence be satisfied that the evidence which may be yielded is relevant and that the recording produced is authentic and original ... The above principles apply to the use of film produced by hidden, automatic security cameras installed in banks and elsewhere for the purpose of recording robberies and other incidents. The jury are entitled to consider the film as identification evidence of the persons recorded on it, subject to the foundational requirements stated above" see eg 'R v Dodson; R v Williams [1984] Crim LR 489; see "Taylor v Chief Constable of Cheshire [1986] 1 WLR 1979'.

It is clear therefore from the text books that Mr Finegan's submission is not correct. They indicate, in our opinion, that the trial judge's perception of what he observed in the film was real evidence and admissible - as much evidence in the case as the sworn testimony or documentary evidence given.

The courts have taken the same stance whenever the modern usage of tape recordings and films in the detection and recording of crime has raised questions of admissibility. First, tape recordings were admitted in certain circumstances, to be put before juries for their hearing and perception of them (see R v Maqsud Ali, R v Ashiq Hussain [1966] 1 QB 688; R v Stevenson [1971] 1 WLR1; R v Robson; R v Harris [1972] 1 WLR 651 and in a number of unreported blackmail cases in this jurisdiction). Then followed a number of cases in which the courts considered the admissibility of video films. One of these was R v Hamilton, Fairey and Noel (unreported, 29 July 1986). He referred to the facts of that case and cited passages from the judgment of Lord Lane CJ:-

> '… we see no reason why video recordings of events such as these should in principle be any more objectionable than the evidence of people viewing the events themselves. The suggestion that the cameramen should have been called we reject. Indeed the video recording has many advantages over the eye witness, because the video recording can be played back as many times

> as you wish. It can be played in slow motion, the picture can be frozen and, as is demonstrated here, stills can be prepared, by methods which are not made plain to us, from the video tape, and we have examples of it before us now.
>
> Similarly, with a video tape the jury can be put into very much the same position as people actually observing the crime being committed, whatever it may be. Of course the weight of the evidence would differ according to the quality of the picture and the way in which the cameramen have done their job'.

Later he said:

> 'The next ground which is put forward with somewhat less enthusiasm is that the learned judge was wrong in directing the jury that if they were not satisfied with the identification given by the police officers, they could use their own eyes and compare the video tapes, both fast motion and slow motion and the stills, with the men who sat before them in the dock, and indeed have sat before them for something like six days, and whose features and appearance they could observe.
>
> The ground of this complaint is to some extent mysterious. We can see no reason why both forms of evidence were not equally admissible and equally cogent: first of all the recognition by the police, if it took place, and secondly, the comparison by the jury of the people appearing in the photographs with the appearance of those who sat in the dock. If the jury were dissatisfied with the police evidence, for instance because they did not like the way in which the video exercises had been carried out, we cannot see any reason why they should not fall back on their own view of the similarity between the defendants and the people who appeared on the photographs if such similarity existed …'.

We consider this case to be strong authority for the legitimacy of the trial judge's viewing of the film (1) for the purpose of identifying and perceiving himself what it showed and (2) for the purpose of deciding the correctness of the police officers'

identifications from it. We respectfully adopt the reasoning and conclusions in the case.

There have been earlier authorities too in which it appears to have been accepted without argument that juries could view video films for their own perception. The contest on admissibility in these cases arose on other grounds. He referred to R v Dodson and Williams [1984] 79 Cr App R 220 and to what Watkins LJ said at p 226:

> 'The submissions made to this court by Mr Mansfield on behalf of Dodson were that whilst he could not argue that these photographs according to the strict rules of evidence were inadmissible, the recorder in the exercise of his discretion should not have done otherwise than to have them excluded seeing that the only purpose of admitting them was to allow the jury to perform the task, by observation of the appellants in the dock and by looking at all the photographs and enlargements provided for them, of identification. This amounted he said to the equivalent or something worse of dock identification which was nowadays not allowed to be done. It took the place of an identification parade which should have been, but which was not, held. Anyone who claimed to identify a defendant as having been present at the scene of a crime should be available - jurors obviously were not - for cross-examination. The jurors were asked to put themselves in the position of eye witnesses without any of the restraints which govern a witness when giving evidence able to be imposed upon them. No one could possibly tell, assuming they convicted, because they relied upon the photographs and enlargements, what in particular led them to that conclusion'.

At p 227, the judgment continued:

> 'Mr Mansfield further conceded that inclusion of the photographs and enlargements could not be properly resisted if the purpose of doing so was to enable a witness, other than a police officer, who knew an appellant well to testify to the effect that having regard to his familiarity with the appellant he recognised him as one of the men in the photographs. The jury could then be left with the

> proper task of assessing the weight of that witness's evidence'.

At p 228, referring to the Crown submissions, Watkins LJ said:

> 'It is, he submitted, in the best interests of the administration of justice that photographs taken at the moment of an attempted robbery be admitted in evidence. They are an invaluable aid in many respects to identification. Argument about clothing worn and weapons carried is almost certainly thereby eliminated and more often than not the identity of the offender clearly revealed ...'

Watkins LJ concluded at p 228:

We entertain no doubt that photographs … are admissible in evidence. They are relevant to the issues as to (a) whether an offence was committed and (b) who committed it. What is relevant is, subject to any rule of exclusion - we know of none which is applicable to this situation - prima facie admissible. As for the exercise of any discretion which a judge has to exclude such evidence in the form of photographs, we have no hesitation in stating that we cannot see any reason why he should do so.

He referred to Taylor v Chief Constable of Cheshire [1986] 1 WLR 1479 and cited what Ralph Gibson LJ said at p 1486:

> 'Where there is a recording, a witness has the opportunity to study again and again what may be a fleeting glimpse of a short incident, and the study may affect greatly both his ability to describe what he saw and his confidence in an identification. When the film or recording is shown to the court, his evidence and the validity of his increased confidence, if he has any, can be assessed in the light of what the court itself can see'"

Kelly LJ went on:-

"Our conclusion on this part of the case is that the trial judge was correct in viewing the heli-tele film both for the purposes of his own recognition of the appellants in it and for the purpose of deciding whether the police officers' recognition of the appellants on it was correct.

Counsel for the appellants then went on to attack the weight given by the trial judge to his own perception of the film.

Mr Cahill submitted he had failed to warn himself of the special need for caution in identification cases before convicting, contrary to the guidelines in R v Turnbull [1977] 2 QB 224. Indeed he went the length of submitting that a passage at p 24 of the judgment suggested that the trial judge considered that Turnbull had no application at all to his own perception of the film although he conceded that the trial judge had applied these Turnbull principles to the police officers' recognition of the appellants on the film. The criticised passage at p 24 reads:

> 'It is relevant to observe that the rules laid down in R v Turnbull and the other rules which have been laid down by the courts in relation to identification were evolved by the courts to guard against the risks of erroneous identification and consequent injustice in cases where the court of trial was itself unable to assess with precision whether the identification made by the witnesses was correct. But where the alleged criminal acts and the events leading up to them have been filmed and have been shown to the court of trial in a video tape the court is in the unusual position, by reason of the advances in science, to see itself the perpetrator or perpetrators of the acts shown on the video film and, if the quality of the video film is good enough, to assess itself whether the identification made by the identification witnesses was correct'.

The guidelines in Turnbull, a decision of the Court of Appeal in England and therefore of high persuasive authority in this jurisdiction, have been accepted and applied here (see R v Maguire [1977] 4 NIJB and R v Russell [1982] 6 NIJB). In Maguire, Lord Lowry referred to 'the wisdom enshrined' in Turnbull and two other like cases, and declared himself aware of 'all the pitfalls' identification cases hold. He added as a further consideration to be weighed, the 'sense of responsibility' of the identifying witness. The continuing importance of the Turnbull guidelines is emphasized in the recently reported opinion of Lord Ackner in Reid (Junior) v The Queen [1990] AC 363 at p376.

We consider that the Turnbull guidelines should be applied and adopted as far as appropriate by a judge in a Diplock court to his assessment of the weight to be given to visual identification made from a video film, whether that identification purports to be made by a witness or witnesses, or by the judge himself. We see nothing in principle to justify a distinction between the consideration of the identification evidence of a bystander and

that of a witness or judge who identifies from a video film screen. The imperfections of human observation, the dangers of suggestibility and the possibilities of honest mistake even by a plurality of witnesses still arise and justify the need for special caution before convicting. In Dodson and Williams (supra) Watkins LJ appears to adapt Turnbull guidelines to the comparing by the jury of what they saw of the robbers filmed by the security cameras with their observation of the accused in the dock. At p 228 he said:

> 'It is, however, imperative that a jury is warned by a judge in summing up of the perils of deciding whether by this means alone or with some form of supporting evidence a defendant has committed the crime alleged. According to the quality of photographs, change of appearance in a defendant and other considerations which may arise in a trial, the jury's task may be rendered difficult or simple in bringing about a decision either in favour or against a defendant. So long as the jury having been brought face to face with these perils are firmly directed that to convict they must be sure that the man in the dock is the man in the photograph, we envisage no injustice arising from this manner of evaluating evidence with the aid of what the jurors' eyes tell them is a fact which they are sure exists.
>
> What are the perils which the jury should be told of? ... We do not think the provision by us of a formula or series of guidelines upon which a direction by a judge upon this matter should always be based would be helpful. Evidence of this kind is relatively novel. What is of the utmost importance with regard to it, it seems to us, is that the quality of the photographs, the extent of the exposure of the facial features of the person photographed, evidence, or the absence of it, of a change in a defendant's appearance and the opportunity a jury has to look at a defendant in the dock and over what period of time are factors, among other matters of relevance in this context in a particular case which the jury must receive guidance upon from the judge when he directs them as to how they should approach the task of resolving this crucial issue'.
>
> In Taylor v Chief Constable of Cheshire [1986] 1 WLR 1479 Ralph Gibson LJ implied and

> McNeill J expressly stated that Turnbull guidelines should apply to identification evidence made from a video recording whether the viewer is a witness describing what he saw or is the tribunal of fact itself."

We are satisfied that if the jury is entitled to engage in this exercise in identification on which expert evidence is admissible, as we have held, there should be a specific warning given to the jurors of the dangers of relying on their own untrained ears, when they do not have the training or equipment of an auditory phonetician or the training or equipment of an acoustic phonetician, in conditions which may be far from ideal, in circumstances in which they are asked to compare the voice of one person, the defendant, with the voice on the tape, in conditions in which they may have been listening to the defendant giving his evidence and concentrating on what he was saying, not comparing it with the voice on the tape at that time and in circumstances in which they may have a subconscious bias because the defendant is in the dock. We do not seek to lay down precise guidelines as to the appropriate warning. Each case will be governed by its own set of circumstances. But the authorities to which we have referred emphasise the need to give a specific warning to the jurors themselves.

In the absence of such a warning to the jury, and there was no such warning given by the Judge to the jurors themselves, we consider that the convictions were unsafe. We are satisfied that if the Court of Appeal had had the benefit of hearing the evidence of Dr Nolan and Dr French on this issue, they would have concluded that a specific warning to the jury themselves was needed.

The third plank on which the case for the prosecution rested was the recognition by Detective Sergeant Johnston of the appellant's voice, based upon hearing a poor quality tape recording of which only 16 seconds were available for voice identification. In view of the difficulties attending voice recognition in such circumstances as explained to us by the experts, we consider that a conviction based on his evidence could not be regarded as safe,

notwithstanding the very suspicious calls made from the injured party's mobile phone after the burglary to the telephone in the house in which the appellant resided some hours after the burglary and the appellant's explanations of them which the jury must have rejected.

Accordingly the convictions will be quashed.

IN HER MAJESTY'S COURT OF APPEAL IN NORTHERN IRELAND

THE QUEEN

v

ANTHONY O'DOHERTY

Appellant.

J U D G M E N T

O F

NICHOLSON LJ
