ELSEVIER

Speech Communication 31 (2000) 205–224

SPEECH COMMUNICATION

www.elsevier.nl/locate/specom

Forensic voice identification in France

Louis-Jean Boë [1]

Institut de la Communication Parlée, UMR 5009 CNRS – INPG – Université, Stendhal, BP 25 38040 Grenoble, France

Received 28 October 1998; received in revised form 19 September 1999

Abstract

Among the sources of information used in legal identification, fingerprints and genetic data seem to provide a high degree of reliability. It is possible to evaluate the probability of confusing two individuals who might possess the same fingerprint characteristics or the same genetic markers, and to quantify the risk of a false alarm. By their very nature, these data do not vary significantly over the course of time, and they cannot be modified by a suspect. The erroneous metaphoric term "voiceprint" leads many people (not only the general public) to believe that the voice is as reliable as the papillary ridges of the fingertips. This is not the case. According to present evidence, certain magistrates in France attach far too much importance to analyses of the voice which, along with other indices, should not be used except to help in directing an investigation. In this communication, the author will detail the conditions under which, in France, voice analyses are carried out in the course of an investigation undertaken by the law, and will attempt to define the limits of this protocol, and the difficulty (and impossibility) of producing a reliable statistical test. A historical review will then be presented of the discussions initiated by and position statements adopted by the French speech community since 1900. Finally some ideas and proposals will be put forward in conclusion, which might be discussed by specialists in speech in collaboration with the police, the *gendarmerie*, and the *magistrature*, on a national, European, and international level, to advance the search for legal proof of identification within a scientific framework, and to end up with well-defined protocols.

Résumé

Parmi les données utilisées en identification juridique les empreintes digitales et les données génétiques semblent présenter un degré de fiabilité élevé. La terminologie métaphorique erronée " d'empreinte vocale " donne à croire (et pas uniquement au grand public) que la voix est tout aussi fiable que les crêtes papillaires des pulpes des doigts. Il n'en n'est rien: un enregistrement de parole n'est pas une trace laissée sur une surface au contact d'une partie du corps d'un individu, ni un prélèvement direct opéré sur celui-ci. Dans certaines enquêtes le seul élément disponible est constitué par un enregistrement d'une communication téléphonique anonyme; on comprend tout l'intérêt que pourraient offrir des techniques fiables d'identification du locuteur. En France il n'existe pas d'experts auprès de tribunaux répertoriés comme spécialistes d'identification vocale. Ce sont des experts en acoustique et vibrations qui sont requis par les magistrats pour des identifications juridiques et les avocats peuvent faire appel à des spécialistes de parole dont le témoignage sera apprécié par les juges. Dans l'état actuel des connaissances il ne semble pas possible de pouvoir procéder à l'identification d'un locuteur. Et pourtant certains "experts" prétendent identifier de façon certaine la voix d'un suspect et certains magistrats accordent, en France, beaucoup d'importance à ces analyses. Actuellement une personne peut être placée en détention préventive pendant plusieurs mois à la suite d'une expertise vocale et elle peut être condamnée sur cette seule preuve. L'auteur précise les conditions dans lesquelles sont faites, en France, les

[1] Rapporteur du Groupe Francophone de la Communication Parlée (GFCP) pour les questions d'identification juridique de la voix.

0167-6393/00/$ - see front matter
PII: S0167-6393(99)00079-5

expertises vocales dans le cadre d'une procédure pénale; tente de cerner les limites de ce protocole, les difficultés (et impossibilités) d'une évaluation probabiliste; présente un rappel historique des discussions et prises de position de la communauté parole française depuis 1990; avance enfin des éléments de réflexion et des propositions qui pourraient être discutées par les spécialistes de parole en collaboration avec la police, la gendarmerie et la magistrature, au niveau national, européen (et international) pour faire avancer la recherche de preuve dans un cadre scientifique et aboutir à des protocoles bien balisés.

1. Introduction: in search of proof of identity

In the past centuries, the identification of slaves, prisoners, mercenaries, soldiers and criminals was one of the many problems facing societies that did not hesitate to brand, scarify, tattoo or handcuff their members. These procedures, which belong to a different age, were still being inflicted not so long ago. To convict or absolve a defendant, courts of law were obliged, for some time, to make do with witnesses' memory of a face or voice. The absence of true parameters for identification sometimes led to irreparable miscarriages of justice.

The need to establish the identity of partners signing a commercial agreement stimulated the use of a reliable, non-invasive, and of course non-traumatic technique for identification: the fingerprint. This has been used by the police for nearly a century; together with genetic fingerprinting techniques discovered and used only recently, the law now has at its disposal truly individual markers that are reliable and stable.

However, in many cases, fingerprints are not available, and the only information available to investigators may consist of a single voice recording, generally made during a telephone conversation. A very pressing and perfectly justifiable demand exists on the part of police and magistrates for establishment of legal proof of identity from measurements of the voice (Fombonne, 1996). The question immediately arises: to what extent can one use a recording of the human voice as the basis for an inquiry, or to establish proof of the guilt or innocence of a suspect? The considerable interest in obtaining reliable techniques for speaker identification, and in using these as the basis for such proof, is easily understood.

In certain cases, the demand is so urgent, and the political pressure so strong that an appeal is made to the general public to identify an anonymous caller. On the 22nd June 1976, for example, a letter written by hand in block capitals and signed *Le groupe 666 des Brigades Rouges* was sent to the *Dauphiné Libéré*, a local newspaper in Grenoble, France. It was a ransom demand. The kidnapper later reiterated his requirements by telephone using public callboxes; he managed to outwit all of the traps put in place by the police, who suspected the work of an extremist group of the kind that currently exists in Italy. The President of France instructed the Minister of the Interior to appear in Grenoble and use all possible means to apprehend the persons responsible. The decision was then taken to broadcast the voice of the kidnapper on the radio and on television. It was the first time that such a procedure had been used in France.

> Scarcely had the voice gone out on the airwaves when telephone calls poured in to the police switchboard. All of the callers believed they had recognized the sound of the voice. (*Le Dauphiné Libéré*, 10 October 1981).

One call in particular alerted the police: a married couple claimed to have recognized the voice, saying that it belonged to the wife's brother, a petty criminal with previous convictions already known to the police. The police, whose inquiry was thus guided by this call among all the others, found incriminating evidence at his home (a typewriter later established to have been the one used to address the messages, and an exercise book containing block letters identical to those in the anonymous letters). Sentenced in 1981, the criminal was found guilty of the murder of the two hostages.

Such a procedure did, in this case, lead to the arrest of the guilty person, but it poses certain risks, as the *Syndicat de la Magistrature* emphasized in denouncing "the use of a mass of informers as a means of resolving a judicial matter of

undeniable gravity", while pointing out that "in the struggle against criminal activity, the ends does not justify the use of any possible means, even in exceptional cases".

2. The legal framework: stages of inquiry

We detail the conditions under which, in France, voice analyses are carried out in the course of an investigation undertaken by the law (Fig. 1).

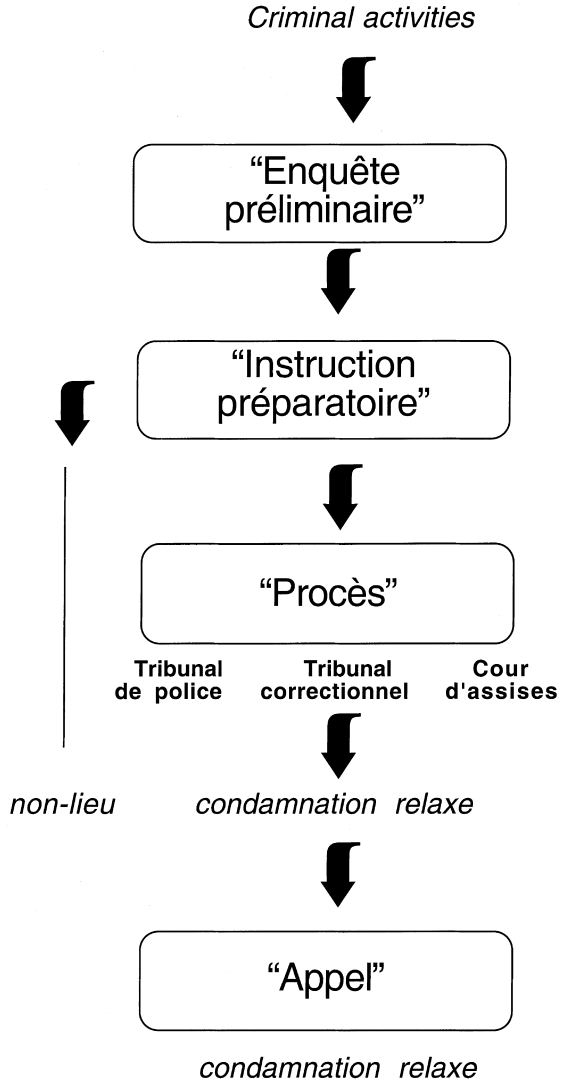

Fig. 1. The stages of the legal framework in France.

2.1. The criminal procedure: society and the individual

Questions of civil law, concerning disputes between individual members of society, will be left aside here, and only the procedures of criminal law will be considered, established by the legislature to protect society, and which require rapid and efficient repression of infractions. Despite this requirement, the legislature imposes a set of limits so that in carrying out this repression the interests of the individual are not sacrificed. If the law determines that a person guilty of an infraction be punished, it also demands that the person under prosecution be afforded every opportunity to defend himself, and preserves the rights of individuals regardless of whether they are considered guilty or not guilty of a crime. The legal process has to permit the prosecution and conviction of the guilty, but must nevertheless prevent the unjust prosecution and conviction of the innocent. "Better to let a hundred guilty persons go unpunished", as the proverb goes, "than to condemn a single innocent man or woman" (Stefani et al., 1996). The penal procedure aims to ensure that punishment is carried out, but a sentence cannot be pronounced unless the guilt and responsibility of the accused have been firmly established by a judge.

By virtue of Article 9 of the Declaration of Human Rights in 1789, the European Convention, and the Constitution of 1958, "any person accused of a crime is presumed innocent until proven guilty in a court of law". As a result of this presumption, a person suspected of committing a crime is not required to prove their innocence. The proof of guilt rests with the prosecution, as it happens with the Ministry of Justice, which has every opportunity for providing this proof, within the limits of the law.

2.2. The stages of investigation

2.2.1. The "enquête préliminaire"

Although in France it is usually the responsibility of the *juge d'instruction* to initiate proceedings based on facts which apparently relate to a crime, the law gives the *police judiciaire* the power

to carry out investigations, where criminal activity is or is not evident, either spontaneously or upon a request from the public prosecutor. Admissible evidence at this stage might consist of fingerprints, traces of blood or sperm, hair, or recordings made using microphones or through phone taps, or by interception of telephone communications (for legal aspects see, for example, (Stefani et al., 1996, pp. 36–37)).

As an example, consider a case where a person informs the police that he is the subject of anonymous threats. They arrange to have his telephone line placed under surveillance and record a new anonymous phone call made from a telephone booth. The *enquête préliminaire* leads to a suspect, all of whose telephone calls are subsequently recorded (with the authorization of the public prosecutor. The police can then carry out a comparison of the anonymous voice and that of the suspect in their laboratories. The report detailing the results of the analysis can be included in the dossier assembled by the *enquête préliminaire*. Once the *police judiciaire* have completed the *enquête préliminaire*, the results are communicated to the public prosecutor, who will make a decision regarding the possibility of a prosecution.

The International Criminal Police Organization, INTERPOL, was created in 1923 to counteract international crime. Currently 177 countries are members, and the new headquarters of the organization was inaugurated in Lyon in 1989. The sections of the penal code that apply to the use of forensic voice analysis are strictly within the jurisdiction of the member states. Article 2 of the INTERPOL statutes specifies that the aims of the organization are notably "to develop and to ensure the greatest possible reciprocal assistance between all police authorities within the framework of the laws existing in the different member countries, in the spirit of the universal declaration of human rights".

2.2.2. The "procédure d'instruction"

When preventive action is required, not all of the proceedings are brought directly before the *juridiction de jugement*. Some of the most complex and serious cases are first of all submitted to a juridiction d'instruction. This is obligatory when a crime has been committed, but optional for a minor offence; it is carried out in secret and does not involve a hearing. The first task of the *juge d'instruction* consists in informing the accused that he has been placed under investigation, and that the evidence merits such a decision. The *juge d'instruction* then proceeds to examine the evidence, and to seek for proof of guilt. To establish such proof, he may make use of written documents, statements by witnesses, or confessions, and he is allowed to carry out searches, seize evidence, or ask for expert advice. In particular, the Ministry of the Interior may intercept and record communications by telephone or FAX. The *juge d'instruction* has the power to organize a judicial hearing, provided that the penalty incurred is greater than or equal to two years imprisonment. The duration of the hearing cannot exceed four months (excepting renewals) and a *procès verbal* must be established for each recording (Chambon, 1991; Maron, Varon, 1991). During the course of the inquiry, the police may not carry out retrospective telephone recordings. The persons recorded may be present at the inquiry, as well as the person under investigation, a representative for the prosecution, and even a lawyer (the president of the Bar must be informed by the magistrate in charge).

Once assembled, the evidence is examined by the *juridictions d'instruction* and according to the nature of the charges they may or may not decide to send the person concerned before *the juridiction de jugement*.

2.2.3. The "juridiction de jugement"

This phase in the different stages of the trial normally follows the *procédure d'instruction*. The courts of common law are not always the same, depending on the offence involved: the *tribunal de police* passes judgement on misdemeanours, the *Cour d'appel* judges infractions and appeals against convictions for misdemeanours, while the *Cour d'assises* judges crimes.

Although the law strictly regulates the search for and processing of evidence, it leaves complete liberty to the judge's examination. The final decision taken is, in effect, governed by the fundamental principle of *intime conviction*. A conviction may equally well be based upon elements assem-

bled during the *enquête préliminaire*, as upon those established during the *procédure d'instruction*, or those appearing in the course of the hearings. The judge has complete freedom to decide upon the value of the evidence which is submitted to him. His decision is made according to his conscience: he condemns, releases or acquits the accused depending on whether he is or is not convinced of his/her guilt, without being obliged to justify the weight he attaches to each of the pieces of evidence available to him. In the case of a misdemeanour (but not for a crime) he must explain the motivation behind his judgement.

2.3. The appeal to expert knowledge

2.3.1. During the "enquête préliminaire"

In the course of the *enquête préliminaire*, a law passed on 30 December 1985 expressly authorises the public prosecutor or an officer of the *police judiciaire* to consult all persons qualified to take part in technical or scientific investigations. The police and the gendarmerie have their own laboratories (e.g. *Laboratoire Central de la Police*, Paris 15eme; *Institut de Recherche Criminelle de Gendarmerie*, Rosny sous Bois). These are capable of carrying out a large number of analyses and examinations in a great many fields: ballistics, analysis of documents (writing, typewriters, printers, toner), typing, fingerprinting, vehicles, acoustics, photography, electronics and computer science, chemistry, biology, explosives, toxicology, anthropology, entomology, legal medicine, odontology, etc. In all matters concerning DNA analysis, the French legal system has begun to consult laboratories in the United Kingdom. In 1985, Pierre Joxe, Minister for the Interior, highlighted forensic science as being among the priorities for his Ministry and five laboratories were equipped.

2.3.2. During the "procédure d'instruction"

In the same way, when the use of technical expertise can lead to the establishment of proof, the *juge d'instruction* can ask for certain points to be clarified by a specialist. This consultation may also be initiated by a request from the prosecution. The procedure of the penal code determines the way in which the scientific consultation is to be carried out. The experts consulted by the *juge d'instruction* are chosen in principle from an official list of experts drawn up by the tribunals. These may consist of individuals or organizations; when organizations are consulted the legal representative of the organization designates the name of the person who will carry out the analysis. Each *cour d'appel* prepares a list of experts practising within its area of jurisdiction. There also exists a national list drawn up by the office of the *bureau de cassation*. The *juge d'instruction* chooses the experts he needs, according to the trust he has in them, either from the national list, or from the lists provided by the *cours d'assises*. The judge may even consult an expert who does not appear on any of these lists, but he must then provide reasons for his choice whenever he is required to explain his final judgement. The judge nominates the expert in a legal ruling and details the investigation that he is to carry out. This investigation must involve nothing but the examination of technical questions. Since 1985, "the *juge d'instruction* has been responsible for designating the expert charged with the execution of the expertise" (paragraph 1). However, paragraph 2 specifies that: "if the circumstances justify it, he may designate several experts". If the public prosecutor requests a different assessment, the *juge d'instruction* can rule that this is not useful and renders a ruling within a month, with an explanation of his decision; if, on the other hand, he considers that it might be useful, he calls upon the expert or experts to testify.

The experts must pledge "to uphold the law on their honour and conscience". They proceed to carry out their assignment under the supervision of the *juge d'instruction*. In principle, the experts cannot interrogate the person accused. The experts are personally responsible for carrying out the operations with which they have been entrusted. The expert is not a substitute for a *juge d'instruction*, but instead must try hard to inform the latter as well as possible on all the points which have been defined to him.

The experts must submit their report within the time limit fixed by the judge, and provision is made in the civil code for the subsequent payment of experts. Following this, the *juge d'instruction* must

assemble the parties involved and their lawyers, and inform them of the conclusions reached by the experts. Within a fixed period of time, all parties may express their opinions and, if necessary, call on counter-experts. The reliability of the expert is not brought into question except in cases of wilful misrepresentation or gross error.

As there does not exist a special expertise in "voice analysis" either in the *Cour de Cassation* or in the tribunals, experts qualified in "acoustics and vibration" can be called upon, or else specialists in speech processing (engineers or phoneticians) or signal processing, or indeed any specialists from other disciplines previously consulted by the police or called upon by the courts. Certain magistrates accord great importance to forensic voice analysis, and a suspect can be detained and then condemned on the evidence of voice analysis alone.

3. Reliable parameters for identification: the fingerprint

3.1. Fingerprints

The first known fingerprints go back to Neolithic times. It seems likely that these identifying marks (thumbprints, for example) may have been used very early on in China, for the authentification of documents. In 1686, the Italian anatomist Marcello Malpighi remarked upon the diversity of human fingerprints. However, it was not until 1823 that J.E. Purkinje published the first treatise presenting a classification of the papillary ridges into 19 different types (the Purkinje drawings), although he does not seem to have foreseen applications in identification. The first of these was due to Sir W. Herschel, an English administrator in India, and was introduced to avoid any possible confusion with contracts and prevent the need for litigation against Bengali contractors. The effectiveness of this method prompted Herschel to pursue the study of fingerprinting. He took measurements from subjects over a period of thirty years and thereby convinced himself of the permanence of the papillary folds. The official use of fingerprints therefore began in India in certain government departments in 1877. It was nevertheless Francis Galton, the famous British geneticist and anthropologist, who carried out the first scientific study from 1888 onwards. In 1894 fingerprinting was adopted in Great Britain as the method of choice for proving identity. Galton statistically evaluated the risk of identification errors as a function of the number of points measured. He compared these evaluations with those he deduced from anthropometric parameters supplied by Bertillon. The Chief of the Metropolitan Police, Edward Richard Henry, demonstrated the incontestable practical validity of the fingerprinting technique.

The skin, which forms itself into ridges, contains the orifices of sweat glands through which sweat, mixed with sebum, is secreted. On contact with a smooth surface, this mixture leaves a recognisable trace which can be revealed by more and more sophisticated techniques. Fingerprints can remain in place for many years on numerous media. Fixed definitively from the age of four months, recoverable from corpses (and mummies), unalterable and irremovable, fingerprints constitute a truly individual characteristic.

At the present time, a simple search on the Internet for the word "fingerprint" yields almost a hundred web-sites, and numerous companies offering commercial products.

3.2. Genetic fingerprints

The continuity of specific characteristics of a species is ensured by information transmitted during cell division. This genetic material, present in each cell, is composed of DNA (Deoxyribo-Nucleic Acid) contained within the cell nucleus, and DNA present in the mitochondria. The molecular structure of DNA was revealed by Watson and Crick in 1953: a double helix formed from two interlaced chains and composed of nucleotides (6×10^9) held together by strong links. Only part of the DNA functions as a code allowing each cell to synthesize the molecules necessary to ensure its operation and regulation. These regions, termed mini-satellites (since they are arranged around the periphery of the DNA), are made up of repetitive sequences of nucleotides (20 to 70) reproduced three to ten times depending on the individual. It

was among these sequences that Jeffreys, Wilson and Thein located, in Great Britain, in 1985, certain families of sequences exhibiting a common central pattern: a molecular chain composed of a sequence of nucleic acid containing of at least 20 nucleotides. As molecules are transmitted by inheritance, it is possible to carry out identity searches on families.

In medical law, molecular chains are used that are not directly implicated in the genetic code, but which contain individual characteristics which enable a line of descent to be determined. Differences and similarities are more easily established from comparison of nuclear DNA, but the tissue samples that are generally available (part of a single hair, bone or tooth) do not contain enough nuclear DNA; however they often do contain a sufficient quantity of mitochondrial DNA (DNA(MT)). In September 1996 in Tennessee, a murderer could be identified thanks to the establishment of a correspondance between the structure of the DNA(MT) in his saliva and in saliva samples that had been recovered on the hair of his victim. To identify human remains, DNA(MT) taken from bone samples is compared with that taken from a person's brothers or sisters, and it is possible to conclude whether the individual belonged to the same family. Since 1993, experts have been positive that the bones of a corpse rediscovered at Yekaterinburg are indeed those of Tsar Nicholas II; an analysis of the ADN(MT) of his family has removed all doubt (cf. Gibbons, 1998).

France, which in the beginning lagged far behind Great Britain, now carries out more than 2000 analyses each year. A law passed on 29 July 1994 specifies that samples of genetic material may not be taken or analysed except in the course of an inquiry or investigation pertaining to the judicial process or for medical aims. Thus in the criminal affair of Pleines-FougËres (the murder of Caroline Dickson) the *juge instructeur* Van Ruymbeke had genetic tests carried out on all male adults in the village less than 35 years of age (169 in total), but the consent of each person was required (only one refused). The actor Yves Montand refused to undergo a paternity test during his lifetime, but on 6 November 1997 the *Cour d'Appel* in Paris ended up ordering his exhumation for genetic analysis; the Ethics Committee pronounced itself "embarrassed" by this decision.

Since the probability of encountering two individuals with the same genetic profile is extremely small (1 in 10^6 for ten bands, and 1 in 10^{12} for 20 bands; Ludes, Mangin, 1992), genetic analysis appears to be the most reliable of the procedures developed to date. The possibilities it offers for establishing parental relationships have enlarged the search domain considerably.

Table 1 presents a number of different comparisons between fingerprints, genetic fingerprints and the speech signal.

4. From the face to the voice

4.1. The face and the "Identikit"

In 1881, Alphonse Bertillon created his famous "Legal Identification" service in Paris. The system to which his name remains attached to this day was, at that period, universally in use. Using photographs of a subject's face and profile, and measures linked to the skeletal bones, Bertillon (1890) devoted himself to an extensive research programme in anthropometry (measuring the distance between the eyes, the dimensions of the skull, the length of the middle finger of the left hand...). The validity of the method rests entirely on two hypotheses (the temporal stability of bone dimensions, and the absence of correlation between different dimensions) and the possibility of obtaining precise measurements. The Bertillon system was replaced after 1900 by the use of fingerprints, which are incomparably more reliable.

Nevertheless, a method of describing the face (the "Identikit") is still in use, which consists of details of morphological particulars (forehead, nose, ear, mouth) and colour parameters (eyes, hair, beard). Using different features, presented in the form of photographic montages, methods of identification using the Identikit are useful in stimulating the memory, improving precision, and producing a document which can in certain cases lead to the arrest of a guilty party.

Table 1
Different comparisons between fingerprints, genetic fingerprints and the speech signal

	Genetic fingerprints	Fingerprints	Speech signal
General characteristics	Cell structure	Traces of the shape of a finger on a surface	Consequences of displacements of body parts (articulators) on air
Parameters	Multilocular parts of DNA	Geometric	Temporal and spectral
Variability	None	None beyond fourth month	Important • can be modified by the speaker (disguise)
Voluntary modifications or modifications induced by situation	None		• modified by conditions of production (stress)
		None	• can be modified by processing and by characteristics of transmission line and recording set
Age influence	None	None	Important with growth
Wealth state influence	None	None	Important in case of vocal & nasal tract and lungs
Language influence	None	None	Very important
Possibility to identify corpse	Yes	Yes	No
Possibility of sex identification	Yes	No	Uncertain
Possibility of family identification	Yes	No	No
Probability of false identification	•$1/4^{10} \approx 1/10^6$ with 10 stripes •$1/4^{20} \approx 1/10^{12}$ with 20 stripes	$1/2^{29}$ with 35 details	Unknown No estimation in the present state of knowledge
Databases	No (in France)	Yes (700.000 in France)	No (for French)
Identification function	Yes	Yes	No
Discriminant parameters	Yes Galton (1892) Cummins, Mildo (1961)	Yes Jeffreys et al (1985)	To be found despite numerous studies
Law conditions	Law of 1994	No constraints	Law of 1994 (France)
Individual consent	Necessary (France)	No Condemnation in case of refusal (France)	No (France)

4.2. The voice and so-called "voiceprints"

4.2.1. The recognition of familiar voices

People possess the ability to recognize familiar voices (Hecker, 1971; Bricker and Pruzansky, 1976; van Lancker et al., 1985a,b): certain brain lesions occurring in the right parietal hemisphere induce the loss of this ability, termed *phonagnosis* (Van Lancker et al., 1988), just as accidents involving the brain may lead to an inability to recognize familiar faces, termed *prosopagnosis*. There exist therefore in the voice certain characteristics which, in certain circumstances, allow the speaker to be recognized. This evidently does not allow us to jump to the conclusion that each voice possesses characteristics which render it unique among all other possible voices.

Aural recognition of familiar voices is, moreover, far from attaining recognition rates of 100%. For example, with 29 known voices, 13 subjects tested achieved a correct identification rate of 31% on the single word "hello", 66% on a whole phrase, and 83% on 30 s of speech (Ladefoged and Ladefoged, 1980).

No proof has been produced that phoneticians, and more generally specialists in speech, are any better in aural identification tasks, even if they have at their disposal the means for carrying out an analysis. For a long time in Great Britain, methods were proposed for establishing the

identity of a speaker on the basis of an auditory judgment. During a trial held in Scotland in 1967, a man accused of being responsible for false fire alarms was declared guilty on account of the testimony of a phonetician requisitioned by the Court. Despite presenting sonagrams to the Court throughout the trial, this phonetician essentially based his testimony on his own auditory impressions. This affair was the subject of extensive media coverage, and evoked a response from the majority of phoneticians. After having put together a working group to define a position on the use of forensic voice analysis in legal proceedings, they made their reservations known to the Home Secretary. Following this, according to Nolan (1983, p. 16), the law courts in Great Britain no longer made use of voice analysis based on sonagrams. At the Colloquium of British Academic Phoneticians in 1980, British phoneticians approved, by 30 votes to 12, a motion stating that: "phoneticians should not consider themselves expert in speaker identification until they have demonstrated themselves to be so" (Nolan, 1983, p. 17). In order to quantify the ability of phoneticians in comparison with naive subjects, a test was carried out by Marion Shirt at the University of Leeds. The phonetic material used consisted of a studio-quality recording of pairs of samples, each of 5 s duration, produced by male subjects. The test was given to phoneticians, who were free to make use of any technical assistance they wished, and non-specialists in speech, using only their innate auditory ability. The mean correct identification rate was 53% for phoneticians, and 46% for non-phoneticians. Specialists equipped with listening facilities (tape loop repeaters) and given all the time they wished to carry out the test, and naive listeners forced to respond to the question "Is it the same voice?" within a limited time period both obtained practically the same results, which incidentally were extremely poor, since they were all wrong half the time and had identification rates that varied widely (38% minimum, 76% maximum for the phoneticians).

It needs to be stated that we are still a long way from having available a complete procedure for drawing up a portrait of the voice: for example a synthesizer which would permit reconstruction of the voice recalled by a witness from a collection of features. This field of research is of course of interest to criminologists. The article published by Yarmey (1991) is of some relevance here, and describes an experiment carried out with 295 psychology students, who listened to 40 different male voice samples, lasting 36 s. The objective was to test the capacity for describing a voice in the days following the original event. For a description of a voice to be reliable, it needs to be carried out by the witness within twenty-four hours of the incident occurring.

4.2.2. A serious terminological error

In 1962 an article appearing in *Nature* entitled "Voiceprint identification" introduced a term which is still in vogue in daily newspapers, televised police dramas, and spy films. This erroneous metaphoric terminology "voiceprint" leads many people (not only the general public) to believe that a graphical representation of the voice (the sonagram, as it happens) is just as reliable as the structure of the papillary ridges of the fingertips, or genetic fingerprints, and that it allows reliable identification of the original speaker. Of course we now know that this is not at all the case: neither Kersta (1962) nor any other specialist in speech has been able to provide evidence of the reliability of speaker identification through analysis of spectrograms. Following a number of controversial judicial testimonies (Tosi, 1979), the Technical Committee of the Acoustical Society of America asked a team of renowned scientists to produce a report. They concluded that:

> Today's consensus suggests that speaker identification by voice pattern is subject to error at a high, and as yet undetermined, rate. Courts' determinations may also depend on the apparent validity of exhibits brought in evidence. Spectrographic evidence may often display features that are overwhelmingly influenced by the words spoken rather than by the speaker's identity. Judge and jury may therefore be misled in understanding the evidence and in assessing an expert's testimony. (Bolt et al., 1970, p. 602).

The use of the expression "voiceprint" is a perversion of terminology (cf. Bimbot and Chollet, 1998) but nevertheless it persists, even, quite remarkably, in reference books.

The relationship between various identification parameters is shown in Fig 2. In the widest sense possible, a recording of speech is not a trace left on a surface in contact with part of an individual's body, nor is it a direct sample taken from the latter: it is in effect nothing more than the indirect record of complex articulatory movements. The speech organs induce variations in instantaneous acoustic pressure which can be recovered by a transducer that converts these into variations in electric voltage. Like all human gestures, speech gestures are not reproducible over the course of time. The parameters used to describe speech clearly show their dependance on the speed of articulation, the loudness of the voice, the psychological state of the speaker, and environmental stress. Automatic speech recognition is in fact directly confronted with this intra-speaker variability, which is intrinsically linked to the processes involved in speech production. In addition, it is evidently important to take into account parameters governing the transmission and recording of

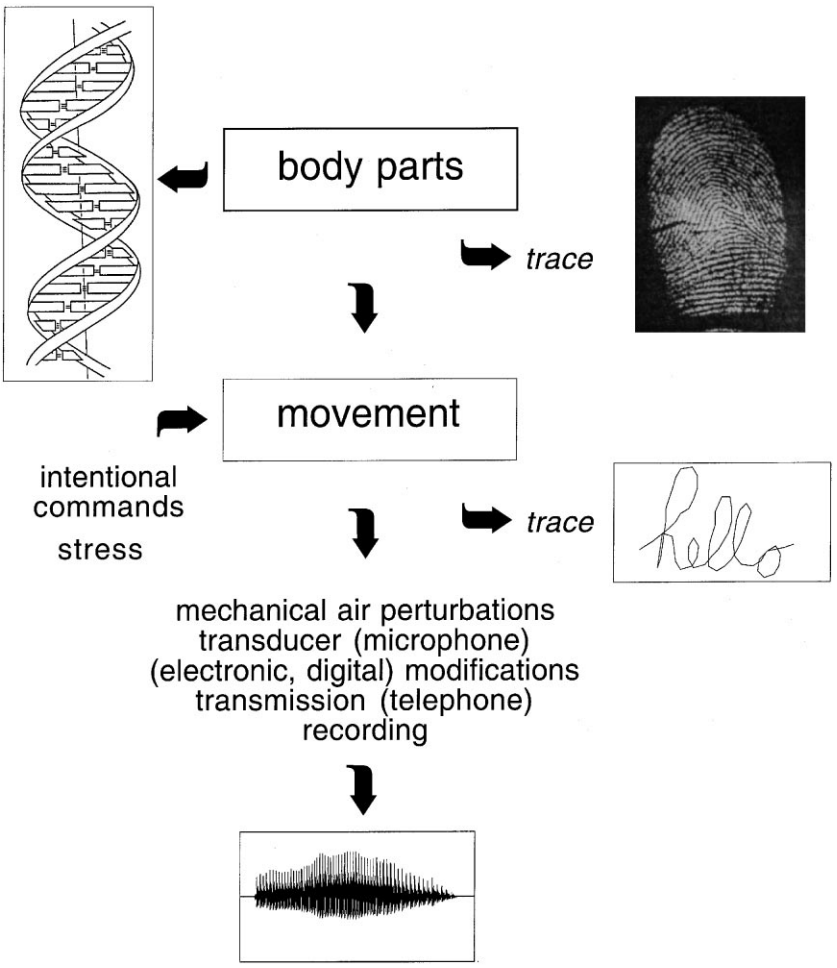

Fig. 2. The erroneous metaphoric term "voiceprint" leads many people (not only the general public) to believe that the voice is as reliable as the papillary ridge of the fingertips. This is not the case. A recording of speech is nothing more than an indirect measurement of complex articulatory movements.

any speech, and the possibility that several voices or noise may have been superimposed. In the case of a recording made over a telephone line, the characteristics of the microphone, the telephone line, and the recorder itself will also enter into the picture. Experts often cannot have access to all these data. To all these factors must evidently be added the possibility of imitation or disguise, and the use of a whole range of techniques for distorting the voice, from a simple spectral equaliser or vocoder to recent techniques such as voice morphing. There even exist digital telephones which provide 16 possible ways of modifying the voice and its loudness. Without knowing the details of the program used, it is impossible to identify even the speaker's gender (Al Bader, 1992).

> A few hours later, Paul Bensoussan received a telephone call from the kidnappers. In a message that was undoubtedly pre-recorded, in an electronically-distorted voice, the gangsters demanded a ransom of 600,000FF [for freeing his wife]. (*Libération*, 15 December 1997).

A message may therefore have been recorded by a speaker in conditions that are inaccessible to the investigators: for example, to avert suspicion, it may have been fabricated by the guilty party from the voice of another speaker.

At the current state of research, no collection of parameters permitting systematic characterization of the voice is known. Could one, however, evaluate the *probability* that two recordings come from the same speaker?

4.3. In search of identification probabilities

How is forensic voice analysis currently conducted in France? To begin with, the police have at their disposal a recording of an anonymous voice obtained (for example) from a phone tap, or carried out by the plaintiff at his or her home (following blackmail, threats, sexual harassment): this is termed the *pièce de question*. A second recording, termed the *pièce de comparaison*, is produced during the *enquête préliminaire* or the *phase d'instruction* by recording the voice of a suspect whom the police ask to read a transcription of the anonymous recording. From these two recordings, a specialist in speech analysis (from the *police judiciaire* or elsewhere), the so-called "expert", must reply to the following question posed by the *juge d'instruction*: do the two recordings come from the same voice, i.e. are the suspect and the anonymous caller one and the same person? Seven possible responses are provided: certain identity, possible identity, probable identity, possible non-identity, probable non-identity, certain non-identity, or impossible to judge.

At this point, it should be remarked that neither the question nor the possible responses include any quantitative statement or justification of the methodology used. There is therefore no way of knowing whether the expert based his response on rigorous mathematical procedures, or on pure intuition. Supposing, hopefully, that some well-defined procedure has indeed been used, current models for speaker identification fall into two classes: deterministic pattern recognition approaches, which rely on establishing a metric on some space of measurable features capable of discriminating between different speakers; and statistical approaches, which attempt to evaluate the probability that a particular speaker produced a particular message, conditioned on any available information. Since the language used in defining the responses available to the expert is essentially statistical (implicit in the use of the qualifiers "possible", "probable"), it will be assumed, for the sake of argument, that a statistical approach is intended.

Any statistical technique for evaluating identification probabilities is necessarily based on an underlying statistical model, or *probability space*, usually denoted as (Ω, E, P) which essentially consists of a set of possible *outcomes* Ω, a collection *events* E, and a *probability measure* P_π possibly described by a set of *parameters* π.

For speaker identification, the set of outcomes Ω can be taken to be the set of all combinations of possible speakers and possible messages (regardless of how we choose to describe these). The set of events F contains subsets of Ω that are chosen to correspond essentially to all collections of speakers and messages that may reasonably be of interest.

The probability measure P_π assigns a prior probability to each possible event in E, and this may be defined either analytically, by constructing some invented model based on expert knowledge, or empirically, by examining sufficient quantities of real data.

Suppose that we denote by ω_i the event containing all outcomes where some particular speaker i definitely produced any message, and by ω_j the event containing all outcomes where some particular message j was definitely produced by any speaker. Let I be the class of events corresponding to all available prior information (the fact that the *pièce de comparaison* was produced by the suspect, etc.). Then, according to Bayes' Rule (Duda and Hart, 1973; Nadler and Smith, 1993), it follows trivially that

$$P_\pi(\omega_i|x_j, I) = \frac{P_\pi(x_j|\omega_i, I)P_\pi(\omega_i|I)}{P_\pi(x_i, I)}.$$

If we choose ω_i to correspond to the suspect, and x_j to correspond to the *pièce de question*, then $P_\pi(\omega_i|x_jI)$ the posterior probability needed to determine the ultimate response to the question posed by the *juge d'instruction*. To calculate this, we need the remaining three quantities in the equation, all of which must take into account the prior information I. $P_\pi(x_j|\omega_i, I)$ is the probability that the message x_j might have, been produced, given that speaker ω_i is talking; in effect, this requires a "production model" for a specific speaker under the circumstances specified by the information in I. $P_\pi(\omega_i|I)$ is the probability that speaker i might have produced any of the possible messages in these circumstances, and could take into account any alibis or auxiliary evidence pointing to this speaker. $P_\pi(x_j|I)$ is the probability that message j might have been produced by any speaker, but since it is the same for all speakers, it does not play any role in the final decision, and can be ignored (Champod and Meuwly, 1998).

Given the size of the population from which the set of possible speakers can be drawn, and the variety of possible messages which might be produced by these speakers, the enormity of the problem of somehow generating a general statistical model for forensic speaker identification is immediately evident. No such model has been produced for French or for any other language, and it cannot therefore be claimed that such a calculation can be carried out, let alone afforded any credibility, at the present time.

Even disregarding the details of the calculation, it is clear from this (albeit over-simplified) example that the final probability, even if this can be established by a logical sequence of mathematical arguments from a well-defined quantitative statistical model, will *always* depend crucially on the initial choice of underlying probability measure. Different probability measures will lead to different probabilities, and hence possibly to different responses to the crude question posed by the *juge d'instruction*. This is true regardless of whether the method used is based on a maximum-likelihood approach, Bayesian methodology, or Neyman–Pearson theory.

Is it possible to obtain a reliable and impartial probability measure which is not subject to manipulation by the expert, or is it possible at least to ascertain how the underlying probability measure has been constructed, and how well it performs? Two basic philosophies are commonly adopted in estimating probability measures. *Non-parametric* estimation techniques rely on using the law of large numbers to guarantee convergence of simple relative frequency estimates of multivariate probability distributions to their underlying true values, but typically only produce reliable estimates when applied to very large amounts of data. Sufficiently large standardized databases of real speech for speaker identification have not currently been adopted or approved by the legal community for this to be feasible, and indeed are not available at the present time. *Parametric* estimation techniques assume that the probability measure can be expressed as a function of meaningful parameters, which are chosen by the expert, and attempt to select these parameters to optimize some sensible estimation criterion for a particular set of training data. When the model assumptions reflect what occurs in the real world, such methods can work well for small amounts of data, but often produce aberrant results if the imposed hypotheses are in error, and may generalize poorly or, conversely, fit particular data sets too closely. Success depends

on the skill of the expert and the form of the model chosen. Again, no standard legal procedure for describing, implementing or evaluating such methods on standardized sets of data exists at present.

The unhappy conclusion to be drawn is that, at present, speaker verification is not at the stage where particular methods can confidently be said to provide reliable quantitative answers to the kind of questions that the jurisdiction is likely to wish to pose. Sufficient quantities of standard data are not currently available to be able to evaluate the performance of existing methods. Worse still, even if perfect models and unlimited data were at our disposition, the legal framework for quantitatively assigning sufficient weight to the reliability of decisions provided by experts has not yet been put in place.

5. Forensic voice analysis and the position statement adopted by the (Francophone) speech communication group

On 16 October 1989, the Ministry of the Interior (Communications and Computer Science Department) launched an appeal for offers for "the study, implementation, and presentation of a means of permitting speaker identification by methods for comparing magnetic recordings. The administration will provide representative samples of the required type. The successful candidate will implement a system providing the best recognition rates possible, will produce a working system, and will carry out the series of tests corresponding to the given samples". The advertisement appeared in the *Bulletin officiel des annonces des marchés publics*. It was communicated to the President of the Speech Communication Group (GCP) of the Société Française d'Acoustique and widely distributed in speech research laboratories.

Taking into account the importance of this scientific problem and the gravity of its social and legal consequences, the members of the GCP committee set out and adopted an official position statement (Appendix A), after consultations with member laboratories. Here are the essential points of the motion, voted unanimously on 7 September 1990; the members of the committee stated:

- that speaker identification is an unsolved problem, and that methods used to date are not reliable;
- that it would be suitable, for ethical reasons, that any specialist professing to be an expert in speaker identification should provide evidence of his competence before providing a forensic voice analysis.

Far from wanting to sidestep the problems posed by the use of voice analysis during a trial or police inquiry, the GCP committee proposed instead that all members should:

- contribute to all research, database collection, experimentation ... which might advance our knowledge about individual characteristics of the voice;
- participate in the establishment of protocols and evaluation of experts, software, or devices that could be put forward to resolve the legal problems of voice identification.

As will be seen later, this motion was reaffirmed in its entirety in 1997.

To the author's knowledge, no laboratory involved in speech research responded to this appeal. *The Institut de la Communication Parlée in Grenoble* took it upon itself to inform the Ministry of the Interior that it was not possible to carry out speaker identification in a reliable manner, and that it would be a very good idea to put in place a number of rigorous evaluation procedures. In collaboration with SAGEM, ICP therefore proposed a project to put together a set of databases that would serve as the basis for these tests.

At least one laboratory replied to the offer and was accepted. This was the *Société Micro Surface* (SMS), an offshoot of the *Laboratoire de Micro Analyse des Surfaces* at Besançon. This laboratory was already known to the Minstry of the Interior for its expertises in ballistics:

> At the beginning of 1990, the Ministry of the Interior made available a large sum of money and launched the offer of a contract for the development of a 'vocal speaker recognition' system. This project was entrusted to the company 'Micro Surface' a small industrial firm

> founded four years ago by researchers from the École Nationale Supérieure de Mécanique et des Microtechniques (ENSMM) at Besançon. The company specialised in the conception and development of instruments for the physical and chemical analysis of surfaces. (*L'Est Républicain*, 19 September 1992).

The motion proposed by the GCP was not widely distributed, but the Ministry of the Interior was aware of it: the *Laboratoire Central de la Préfecture de Police de Paris* (LCCP) organized *a "Journée d'échange sur la comparaison de locuteurs"* on its premises on 22 November 1992. It brought together 33 participants from the French and Belgian police, the *École Nationale de la Magistrature*, two legal experts, and two representatives of the *Société Micro Surface* from Besançon (in charge of the project arising from the offer); the speech community was broadly represented, and several members of the GCP committee were present.

At the start, the representative for the magistrature stated all of the misgivings that this type of analysis aroused: the absence of official experts registered in this special field; the reliability of methods which did not seem to him to have been proved; the lack of certainty in any conclusions drawn.

He drew attention to the questions posed by the comparative study of: the suspect recording, obtained in unfavourable technical conditions and in a state of emotion or fear, which alters the voice. the comparison recording, recorded in much better technical conditions, but with more or less cooperation on the part of the suspect, who might try to modify his or her voice.

Several scientific presentations addressed the technical problems involved in speaker identification, and the motion of the GCP was read out and commented upon. Lecturers from ENST specialising in signal processing argued at length on the impossibility of arriving at a reliable method of identification. They drew attention to the fact that it is possible to modify the voice using very simple devices.

Speech experts who had been consulted previously on legal matters mentioned the same intrinsic difficulties in speaker identification. Not one of the participants claimed that it was possible to tell with complete certainty whether two recorded voices did or did not come from the same speaker. Mlle. Dalloul Wehbi of the *Société Micro Surface* stated that she shared this opinion, and announced that her laboratory was searching for new invariants able to identify a speaker. She referred to the use of the long-term spectrum (LTS), but specialists in speech expressed their reservations about this technique, based on their previous experience of its poor discriminative ability and poor robustness under different transmission conditions. Two months later, Mlle. Wehbi asked to be nominated as an expert in speech in a criminal case that received widespread coverage in the media, which led *Avocat Général* to declare that:

> I have witnessed many battles in court between experts, but never before the spectacle of an expert coming and begging to be called upon to give evidence. It makes a mockery of the law! (*L'Est Républicain*, 18 November 1993).

6. Conflicting examples of the criminal procedure

Certain aspects of two criminal cases will be presented here: in the first, a court of law refused to call upon forensic voice analysis which appeared to lack reliability, whereas in the second, forensic voice analysis was the determining factor.

6.1. A "revolutionary technique"

Thanks to the contract signed with the Ministry of the Interior, the *Société Micro Surface* developed a software package called REVAO (*Reconnaissance Vocale Assistée par Ordinateur*).

> The invention of the program mobilised the resources of an entire team of people: Mlle. Dalloul Wehbi, doctor of physical sciences and manager of 'Micro Surface', M. Claude Roques-Carmes, laboratory director [...]. Their revolutionary prototype is capable of

'taking apart' all of the signals of a voice. 'Our goal is to find a single invariant parameter that characterizes each individual', commented Mlle. Wehbi. 'We hope to find a voiceprint, similar to the fingerprint that exists already'. (*L'Est Républicain*, 19 September 1992).

Maître Joël Lagrange learnt of the existence of this program, which was presented as being revolutionary. He happened to be one of the three defence attorneys for Albert and Monique Villemin, paternal grandparents of little Gregory, murdered on 16 October 1984, in the trial of their daughter-in-law Christine, the mother of the victim. On 18 September 1992, a month before the meeting organized by the LCPP in Paris, the AFP issued the following notice:

Me. Lagrange has stated to the AFP that a micro-computer has just been produced, which reconstructs the voice spectrum of any person as a colour image. According to him, the machine allows one to tell with absolute certainty whether two voices are distinct, and to establish with 95% certainty whether two voices are those of one and the same person. The public prosecutor thus intends to have the voices of the protagonists in the affair compared with those obtained by recording the telephone calls of the informer, in order to absolve or unmask the anonymous correspondant who announced the death of little Gregory and the place where his body was found. (*Agence AFP*, 18 September 1992).

Me. Lagrange's request was submitted to the *Chambre d'Accusation* in Dijon at the end of 1992, and was rejected by the magistrates. The following year, during the trial of Mme Villemin, the presiding judge, decided that the court ought to have some idea of the reliability of this ''discovery'' before calling for an analysis. He asked Professeur C. Roques-Carmes, scientific advisor to the *Société Micro Surface*, to testify, as well as the manager of SMS, Mlle. Wehbi, who asked to be called as an expert in speaker identification:

Unfortunately the two specialists do not agree. According to M. Claude Roques-Carmes: 'A voiceprint does not yet exist in the sense in which one speaks today of fingerprints or genetic fingerprints. Perhaps this will happen some day, but for the moment we are not yet there. There is no 100% certainty of identification, except if the document provided is recent and of exceptional quality'. This is evidently not the case for cassettes recorded ten years ago by the victims of the anonymous caller on old tape recorders lent to them by the police at Corcieux. Madame Dalloul Wehbi, on the other hand, fights tooth and nail to defend what she considers to be 'her' invention, and which she alone claims to know how to operate correctly. [...] She would have to hear the cassettes before making a decision. The tape recordings are listened to again. The most that can be said is that they aren't of high quality. Nevertheless, Mme. Wehbi is happy enough to say that 'In my opinion, I can guarantee the result 100% if I can be provided with voices to compare.' [...] At the end of the hearing the presiding magistrate Ruyssen clearly asks the question: 'Supposing you are given a cassette of the anonymous caller's voice, and cassettes containing voices to be compared, can you tell us with 100% certainty which is the voice of the anonymous caller?' Both of the two specialists stick to their previous positions: 'No', replies M. Roque-Carmes, 'the reliability of the method cannot be guaranteed; there is too much noise in the background', whereas Mlle. Wehbi, who is neither a phonetician, nor an acoustician, but who entered the field with a different outlook, claims the contrary, on condition, she repeats, that the references provided are reliable. (*Le Bien Public*, 18 November 1993).

Following these contradictory testimonies, the presiding magistrate at the Cours d'Assises decided not to have a voice analysis carried out.

The engineers at *Micro Surface*, who specialise in the analysis of surfaces, had begun by using processing techniques that work well in that field: fractal and multi-fractal analysis. However, since

disappointing results had been obtained, they turned their attention towards analysis of the long-term spectrum (LTS), which since 1961 has been well-known to specialists in speech (cf. Boë et al., 1984) who evaluated the performance of this parameter for characterizing speakers, and found it gave poor discriminability.

The work carried out by the *Société Micro Surface*, in the context of the contract with the Ministry of the Interior, did not produce a single publication, either in national or international journals. The software package REVAO was never commercialised. At the same time as the Villemin trial was taking place, the *Société Micro Surface* was already in liquidation and its assets had been sequestered.

6.2. A possible miscarriage of justice

The following affair was related in the press:

> Jérôme Prieto is Basque [...] on 7 September 1996, an unknown person used a public telephone booth close to Prieto's home to claim responsability for the destruction of several vehicles belonging to policemen. The message was listened to. Due to his geographical proximity and political profile, Prieto was placed under surveillance. A police report [of Mlle Dalloul Wehbi] then 'formally' established the link between Prieto's voice [in French] and that obtained from the telephone call [in Basque]. (*Libération*, 6 November 1997).

J. Prieto was taken in for questioning at Bayonne on 26 February 1997, tried, and imprisoned in Paris.

> In the middle of the month of March, to establish a new comparison, the trade unionist was asked to read a transcription of the text of the telephone call. It was in Basque. [...] A counter-expertise expedited by the gendarmerie itself also concluded that the voices were the same. (*Libération*, 6 November 1997).

In 1997, at the end of the year, the president of the speech group, *le Groupe Francophone de la Communication Parlée* (GFCP), was contacted by one of the attorneys for the defendant, who had been put on trial for acts of terrorism and placed under arrest following the results of an analysis of recordings made over the telephone during the course of the *enquête préliminaire* and a counter-analysis requested by the lawyer at the *cours d'instruction*. The author was asked by the president of the GFCP to contact the lawyer, and prepared a detailed report concerning the position of the GFCP regarding the use of voice analysis, the history of the motion passed on 7 September 1990, and the content of the meeting organized by the GCP on 22 November 1992. At a meeting on 10 October 1997, this report was unanimously approved by the committee of the GFCP, who felt that the motion adopted 7 September 1990 needed no further updating. The motion was thus passed in its entirety once again, and the committee took it upon itself to reaffirm that there did not exist at that time any scientific method permitting two voices to be formally concluded to be identical. The entire file (containing the report and motion) was forwarded to Prieto's defence attorney.

After this testimony, the accused was released on bail. He spent almost 10 months in detention. During his trial, the president of the GFCP and the author of the report were heard as witnesses; They drew attention to the committee's position and scientific opinion, notably the current impossibility of carrying out reliable forensic voice identification. The court condemned J. Prieto to three years in prison, following the recommendation of the *Procureur*, who refused to call into question the scientific validity of voice identification techniques. His lawyer appealed. Recalling the evidence presented during the preliminary investigation and the testimony of the representative of the GFCF and its president, and having heard the President of the *Société Française d'Acoustique* confirm the unreliability of forensic voice identification, the Court of Appeal in Paris considered that:

> The forensic voice identification of J. Prieto [...] is only a single indication of guilt that does not, in the absence of other corroborating elements of the inquiry or preliminary

investigation, constitute a sufficient weight of evidence against the accused, who should be given the benefit of the doubt.

In consequence, partially overturning the deferred judgement, the court declared Jerome Prieto not guilty of the deeds of which he was accused, and released him from prosecution.

7. Conclusions

Is it possible to confirm scientifically that two voices come from the same **speaker with complete certainty**, given recordings carried out under the best possible technical conditions, and corresponding to the same text? The answer to this crucial question is, at present, **no**. Of course, it is always possible to calculate distances or probabilities that quantitatively enumerate differences or similarities for a given parameter or set of parameters, but no scientific consensus exists on the choice of parameters permitting distances to be evaluated, or on the choice of prior probability distribution required to establish a statistical model.

Even if there were a positive response to the question thus posed, there is nothing to suggest that this would remain valid for recording conditions that are often very different from a technical viewpoint, and for production conditions that are often very different from a psychological point of view.

Since 1980, the date on which the first request was presented to them, so-called experts who answer this crucial question in the affirmative, and with complete certainty, have not yet provided the proof that they obtain better results than naive subjects who are simply asked to listen to two recordings and reply whether or not they come from the same speaker.

Given that all of the evidence assembled during the course of the *enquête préliminaire* and *the cours d'instruction*, as well as all of the evidence appearing in the course of the trial, can be taken into account by the judge – who has complete freedom to use this evidence as he pleases to decide whether or not the accused is guilty – one can understand the degree of importance that the result of a forensic voice analysis, establishing with certainty the identity of an anonymous caller, might have.

8. Proposals

In a legal context, speaker identification poses very real problems of professional ethics. Can a specialist in speech analysis provide a court of law with answers to crucial questions that have heavy penal consequences, when those answers have no attested scientific validity? Can an expert claim to possess scientific competence in a domain, without providing any proof of this competence? Can one "beg" to be appointed as an expert in voice analysis, while at the same time claiming to be the only one able to use an invention supposedly capable of identifying the voice?

Our field of research does not possess any formal structure or professional organization, unlike those of doctors, lawyers or architects, nor even a code of ethics, as subscribed to by journalists. The point is not to issue orders, but rather to develop scientific and ethical position statements, to make sure they are known, and to encourage their adoption by all specialists in speech (Boë, 1998). In France, the GFCP provides a framework that is well-adapted to the kind of reflections that are needed, and to implementation of the initiatives that should be launched and applied, like those established by Working Group of the PRC *Communication Homme-Machine* during the ESCA Workshop on Automatic Speaker Recognition and Identification (1994). The following suggestions may be useful:

First of all, and most importantly, experts called upon to testify in a trial should be required to provide proof of their competence.

Secondly, it would be desirable for the *Garde des Sceaux* to be kept informed of all position statements issued, and for the 33 *Cours d'Appel* and the 181 *Tribunaux de Grande Instance* to be warned of difficulties in voice analysis.

Thirdly, it is important to continue the dialogue with the magistrature, so that the establishment of legal proof can be continually improved. In particular, assurances must be obtained from the

instructing magistrates that any expert will be provided with several subject recordings, and several recordings for comparison, so that a blind test can thereby be used to test the expert.

Fourthly, contacts with the police that have already been initiated must be maintained, while in the meantime staying firm and vigilant on the position statements defined by the GFCP. Databases, protocols, and evaluation procedures for speaker identification must be put in place, inspired by work already carried out in speech recognition.

Finally, on a European level, we must strive towards the establishment of a working group which could aim to pass information between the different speech communities regarding legal and police procedures for each country. The American experience, which is well in advance of our own (cf. for example, (Bolt et al., 1970)) and the position statements worked out by the International Association for Forensic Phonetics, should be taken into account. A comprehensive position statement, taking into account the specific details of different nationalities, could be developed, for example by an ESCA commission, and adopted and promulgated throughout the European Community.

The misinformation propagated about speech through the media is often distressing: all specialists in speech are partly to blame. "Voiceprints" and "revolutionary technologies" for voice identification will inexorably pursue their career in the media unless we take steps to demythologise them now.

Acknowledgements

Many thanks for information criticisms, details and discussions to: C. Bernasse-Chevrier (Chef du Service de Documentation Générale, Interpol, Lyon), Dominique Bertrand d'Aléchamps (doctor, Grenoble), Serge Caillet (Directeur Adjoint de l'Institut de Recherche Criminelle de la Gendarmerie, Rosny sous Bois), Jean-Louis Croizier (École Nationale de la Magistrature, Bordeaux), Jacques Fombonne (Conseil d'État, Paris), Elizabeth Lhote (Laboratoire de Phonétique, Besançon), Yolanda Molina Ugarte (Avocate, Bayonne), Emmanuel Mazer (Leibniz, Grenoble), Claude Roques-Carme (Laboratoire de Micro Analyse des Surfaces, Besançon), Michel Rumeau (Laboratoire Central de la Préfecture de Police, Paris), Gaston Tuaillon (Centre de Dialectologie, Grenoble), Jacques Peyraud (Société Nationale de Programme France Région 3, Paris), the records offices of the newspapers *Le Bien Public* (Dijon), *Le Dauphiné* (Grenoble, Veurey), *L'Est Républicain* (Nancy), the members of the GFCP committee, and in particular its President Frédéric Bimbot, Jean-François Bonastre and Pascal Perrier, and at ICP, Christian Abry, Christian Benoît, Rafael Laboissière, Jean-Luc Schwartz, Dominique Vuillet, and Hélène Mathéry.

Thanks to Gordon Ramsay for translation.

Appendix A

The motion adopted by the board of the Speech Group, Acoustical Society of France, 7 September 1990 and voted again in 1997.

Motion adopted the 7th September 1990 by the Board of the *Groupe Communication Parlée* of the *Société Française d'Acoustique*.

Speaker identification poses problems that are both ethical and scientific in nature. It seems to us that its technical aspects can only be studied by research laboratories or industrial concerns with the specific skills and means appropriate to the study of speech.

In the current state of the art, it does not seem possible to identify speakers by automatic methods based on procedures designed by an expert and using a speech sample recorded in optimal (i.e., laboratory) conditions – and even less so if the speech sample is of short duration (e.g., around twenty seconds), has a significant amount of background noise, was recorded in poor technical conditions and/or over a telephone, or comes from a speaker disguising or artificially modifying his or her voice.

The *Bureau du Groupe Communication Parlée (GCP) de la Société Française d'Acoustique* states that the identification of an individual by his or her voice is, to the best of its knowledge, a problem with no current solution. Studies on the subject

published up to the present show the unreliability both of direct auditory expertise and of visual or automatic examination of voice analysis products (spectrograms, etc.). The limited results obtained in laboratory situations are naturally worsened in real situations when the speaker is to be identified from a short recording with background noise (often made over the telephone), even supposing that the speaker has not disguised his or her voice.

The members of the GCP have worked and continue to work on the problem of individual characteristics of speakers, including the question of automatic speaker recognition. As scientists, they are ready to apply their knowledge to any research project which may advance understanding in this field, even if merely by proposing rigorous procedures for evaluating speaker identification systems.

However, their experience allows them to fully appreciate the lack of reliability of the solutions to this problem as proposed up until now. Given the obvious social and moral consequences of rash claims in this field, the GCP Bureau affirms that, in its opinion, speaker identification experts have yet to furnish any verifiable proof of their abilities. In particular, because of ethical concerns, it is incumbent upon any specialist to demonstrate his or her competence in speaker identification before assuming the authority of or operating as an expert (e.g., in law-enforcement or judicial affairs).

Far from wishing to avoid this problem, the GCP recognizes the fundamental need to define and specify, in optimal conditions, the current limits of scientific techniques. For the sake of progress in this field of research, and within the framework of their respective laboratories, the members of the GCP Bureau, if solicited, are willing: to contribute to any research, database collection, experimentation, etc., with the goal of increasing understanding of individual characteristics, and to participate in the elaboration of protocols and in the evaluation of experts, materials, or software when these are presented, in order to resolve the problem of speaker identification in a judicial context.

At its last Board meeting, held in Paris on October 10th 1997, the Board of the GFCP unanimously adopted the following statements:

- The GFCP Board considers that the motion adopted the 7th September 1990 by the Board of the GCP on the problems raised by speaker identification remains entirely up-to-date. This motion is integrally voted again.
- In addition, the GFCP Board reaffirms that, at the time being, no scientific method exists that could lead to the formal identity of two voices.

References

Al Bader, H.O., 1992. The problem of voice disguise. In: 10e Colloque International de Police Scientifique. Lyon, 24–27 November.

Bimbot, F., Chollet, G., 1998. In: Gibbon, H.D., Moore, R., Winski, R. Mouton de Gruyter (Eds.), Handbook of Standards and Ressources for Spoken Language Systems.

Boë, L.-J., 1998. L'identification juridique de la voix: le cas français. In: Historique, problématique et propositions. RLA2C Speaker Recognition and its Commercial and Forensic Applications. Avignon, France, pp. 222–239.

Boë, L.J., Pascal, D., Raymond, D., 1984. Les caractéristiques individuelles de la voix. Bibliographie indexée [1261 titres]. Note Technique NT/LAA/TSS/226. CNET, Lannion.

Bolt, R., Cooper, F.S., David Jr, E.E., Denes, P.B., Pickett, J.M., Stevens, K.N., 1970. Speaker identification by speech spectrograms: A scientists' view of its reliability for legal purposes. Journal of the Acoustical Society of America 47 2 (2), 597–612.

Bricker, P.D., Pruzansky, S., 1976. Speaker recognition. In: N.J. Lass (Ed.), Contemporary Issues in Experimental Phonetics. Academic Press, New York, pp. 295–326.

Champod, C., Meuwly, D., 1998. The inference of identity in forensic speaker recognition. In: RLA2C Speaker Recognition and its Commercial and Forensic Applications, Avignon, France, pp. 125–133.

Duda, R.O, Hart, P.E., 1973. Pattern Classification and Scene Analysis. Wiley, Chichester.

Fombonne, J. 1996., La Criminalistique. Que Sais-je, PUF, Paris.

Galton, F., 1892. Finger Prints. Macmillan, London (Reed. 1965, Da Capo Press, New York).

Hecker, M., 1971. Speaker recognition: An interpretive survey of litterature. ASHA monographs, 16.

Jeffreys, A.J., Wilson, V., Thein, S.L., 1985. Individual-specific fingerprints of human DNA. Nature 316, 76–79.

Kersta, L.G., 1962. Voiceprint identification. Nature 196, 1253–1257.

Nadler, M., Smith, E.P., 1993. Pattern Recognition Engineering. Wiley, Chichester.

Nolan, F., 1983. The Phonetic Bases of Speaker Recognition. Cambridge University Press, Cambridge, UK.

Stefani, G., Levasseur, G., Bouloc, B., 1996. Procédure Pénale, l6e Edition. Dalloz, Paris.

Tosi, O., 1979. Voice Identification. Theory and Legal Applications. University Park Press, Baltimore.

van Lancker, D., Kreiman, J., Emmory, K., 1985a. Familiar voice recognition: pattern and parameters. Part I: Recognition of backward voices. Journal of Phonetics 13, 13–38.

van Lancker, D., Kreiman, J., Emmory, K., 1985b. Familiar voice recognition: pattern and parameters. Part II: Recognition of rate-altered voices. Journal of Phonetics 13, 39–52.

van Lancker, D.R., Cummings, J.L., Kreiman, J., Dobkin, B.H., 1988. Phonagnosia: A dissociation between familiar and unfamiliar voices. Cortex 88 (24), 195–209.

Yarmey, A.D., 1991. Description of distinctive and non-distinctive voices. Journal of the Forensic Science Society 31, 421–428.