

1989
ID
4

Ecole nationale supérieure de bibliothèques

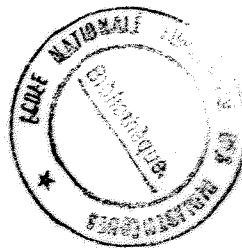
Université Claude Bernard - Lyon I

Diplôme d'études supérieures spécialisées

Informatique documentaire

Note de synthèse

**L'identification du locuteur dans
un but légal**



présenté par **Philippe BERATO**

sous la direction de **Pierre LAVOREL (†)**

1989

1988 - 1989

ID
4

Ecole nationale supérieure de bibliothèques

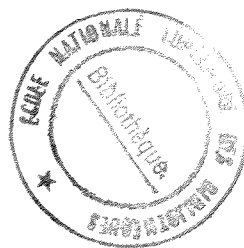
Université Claude Bernard - Lyon I

Diplôme d'études supérieures spécialisées

Informatique documentaire

Note de synthèse

**L'identification du locuteur dans
un but légal**



présenté par Philippe BERATO

sous la direction de Pierre LAVOREL (†)

1989

1988 - 1989

ID

H

SOMMAIRE

| | |
|----------------------------------------------------------------|-----------|
| I Introduction | 4 |
| II Méthodologie | 8 |
| 1 La compréhension du sujet | 9 |
| 2 Les étapes de la recherche | 10 |
| A) Recherche manuelle | 10 |
| a) Le centre de recherches linguistiques et sémiologiques | 11 |
| b) Bibliothèque interuniversitaire, section médecine (Lyon) | 11 |
| c) Bibliothèque interuniversitaire, section lettres | 12 |
| d) Bibliothèque interuniversitaire, section sciences | 12 |

| | |
|----------------------------------------------------------|-----------|
| e) Institut de phonétique de Grenoble | 13 |
| f) Ecole nationale supérieure de police | 14 |
| g) La mise à disposition des documents | 15 |
| B) Recherche automatisée | 16 |
| 3 Conclusion | 16 |
| | |
| III Synthèse | 17 |
| | |
| 1 La méthode des sonagrammes | 19 |
| | |
| A) Les sonagrammes du professeur Kersta | 19 |
| a) Les recherches | 19 |
| b) Critiques | 22 |
| α) Apports de J. P. Mamoux | 22 |
| β) Remarques de P. Corsi | 24 |
| | |
| B) Les sonagrammes du Docteur Anghelescu | 25 |
| a) Les recherches | 25 |
| α) Remarques sur l'appareil phonorespiratoire | 25 |
| β) La méthode d'expertise | 27 |

| | |
|-----------------------------------------------------|-----------|
| b) Critiques | 29 |
| 2 L'identification de la voix par ordinateur | 30 |
| 3 Conclusion | 33 |
| IV Bibliographie | 34 |

I Introduction

Le langage parlé naturellement présente un certain nombre de caractéristiques qui en rendent le traitement très complexe.

L'absence de silence entre les mots pose le problème très délicat de la segmentation d'un flot continu de mots enchaînés.

La très grande variabilité que peut présenter un même discours selon la façon de parler, du locuteur qui peut chanter, crier, murmurer, être enroué ou enrhumé, et aussi selon le locuteur lui-même (homme, femme, enfant, voix nasillarde, voix de tête, différences de timbres, sans parler des accents régionaux, rend très délicate la définition d'invariants.

Les perturbations apportées par le microphone (selon le type, la distance, l'orientation) et l'environnement (bruit, réverbération) compliquent encore le problème.

Par ailleurs; le fait que la parole résulte d'une combinaison d'informations de caractère très différent, se situant à plusieurs niveaux, multiplie les critères de décision.

Ainsi, dans la parole se situent les niveaux :

- phonétique : les sons élémentaires utilisés par une langue.*
- phonologique : la manière dont ils influent les uns sur les autres dans un mot ou une phrase.*
- syntaxique : la structure de la phrase.*
- sémantique : le sens du discours.*
- pragmatique : l'adéquation au contexte réel.*

Enfin, les chercheurs sont confrontés à la nécessité de faire appel à des connaissances relevant de domaines très différents : physiologie, psychologie, linguistique, acoustique, informatique, traitement du signal et de traiter un grand volume de données pour dégager les informations pertinentes à la reconnaissance de la parole.

En dépit de tous ces problèmes, depuis un demi-siècle, des chercheurs tentent de concevoir des dispositifs permettant l'interaction vocale entre l'homme et la machine. En effet, la reconnaissance de la parole et, plus particulièrement, la reconnaissance du locuteur, intéressent de nombreux organismes attirés par leurs perspectives d'application.

Elles peuvent permettre :

- le contrôle d'accès à des processus physiques dans des zones contrôlées (bases militaires, coffres-forts).

- l'accès privilégié à des informations réservées (centres de renseignements, banques de données médicales, bancaires, renseignements généraux).

- la validation des transactions et opérations bancaires (guichets externes : transactions "à domicile", guichets internes).

- une plus grande sécurité pour la vente par correspondance, le client étant authentifié par une "signature vocale".

- l'affectation de tâches nominatives et confidentielles (utilisation d'appareils ne devant servir qu'à une seule personne).

- des applications domestiques : protection de domiciles, garages, voitures par verrou électronique.

- l'amélioration des téléconférences avec une possibilité d'accès réservé aux membres de la réunion.

- la détermination de l'état émotionnel du locuteur : applications criminologiques, mais aussi moyen de juger de l'efficacité d'une méthode pédagogique en enseignement assisté par ordinateur (adaptation au sujet en fonction de son ennui, son intérêt...)

- la recherche de sujets en criminologie.

Seul ce dernier aspect intéresse notre projet de recherche : l'identification d'un locuteur dans un but légal.

L'intérêt de cette reconnaissance est de plus en plus considérable. Le téléphone a été adopté comme "outil de travail" par nombre de malfaiteurs pour préparer leurs méfaits ou les accomplir. Dans les affaires d'enlèvement, les prises d'otage, les rackets, l'identification de la voix peut être vitale. Mais dans de nombreux autres cas moins dramatiques, tels les appels de mauvais plaisants, les fausses alertes à la bombe, la possibilité d'identifier le délinquant serait un facteur dissuasif.

II Méthodologie

1 - La compréhension du sujet :

Une recherche sur la reconnaissance de la parole fait appel à des connaissances relevant de domaines très différents : physiologie, psychologie, linguistique, acoustique, phonétique, physique, informatique, traitement du signal...

D'une part, comme l'a analysé Georges Devereux dans "De l'angoisse à la méthode", l'appréhension d'un nouveau domaine ne se fait jamais sur un terrain vierge, elle se heurte aux acquis culturels de celui qui s'y prête, à ses préjugés.

D'autre part, si le dialogue documentaliste-chercheur est biaisé par les présupposés du premier, il l'est également par les évidences du second. Celui-ci baigne tellement dans son travail qu'il n'arrive pas à se mettre au niveau de ceux qui y sont étrangers.

La présente recherche documentaire n'a pas échappé à ces écueils qui, pour être ordinaires, n'en sont pas moins perturbants.

Ainsi, il a fallu trois rencontres pour arriver à déterminer exactement le sujet. Durant la première, j'ai filtré le discours de M. Lavorel en ayant l'idée que la recherche qu'il me proposait concernait plutôt le domaine informatique et linguistique. J'ai donc commencé par une visite à la bibliothèque du Centre de recherches linguistiques et sémiologiques de Lyon II (CLRS).

Deux autres entrevues furent nécessaires avec mon directeur de mémoire. Celui-ci était à la fois intéressé par mes efforts de délimitation et l'ensemble des pistes que je lui procurais, en tant que sujets proches de son champ disciplinaire.

Une visite à l'institut de phonétique de Saint Martin d'Hères (Grenoble III) fut déterminante. Les références que j'en ramenai permirent à M. Lavorel de préciser ce qu'il attendait exactement de moi : une recherche bibliographique sur l'identification du locuteur dans un but légal

Le but de son propre travail était d'arriver à déterminer des critères certains de reconnaissance pour identifier une voix enregistrée au téléphone et de pouvoir affirmer son origine, avec autant de certitude que dans le cas de la reconnaissance d'empreintes digitales.

Au cours de cette élaboration, j'ai cherché à avoir une idée globale du cadre dans lequel se situait la reconnaissance de la parole. Pour cela, j'ai fait une revue de sommaires de quelques périodiques de vulgarisation scientifique, parus au cours des cinq dernières années :

**Sciences et vie
Sciences et avenir
La recherche**

... pour en lire les articles sur la question.

2) Les étapes de la recherche :

A) Recherche manuelle :

Elle a été effectuée dans différents lieux au fur et à mesure de l'évolution du travail.

a) Le centre de recherches linguistiques et sémiologiques :

Ce centre, inclus dans l'université Lyon II, possède une bibliothèque. Dans le fichier matières, à l'entrée :

**Acoustique
Phonétique
Traitement du signal et de la parole sur ordinateur**

... se trouvent une trentaine de références. Si celles-ci se rapportent bien au sujet, elles sont trop générales et je ne les ai pas utilisées.

b) Bibliothèque interuniversitaire, section médecine (Lyon) :

Le sujet touchant à la neuro-psychologie, j'ai effectué une recherche dans des bibliographies courantes et rétrospectives recensant et analysant les articles de périodiques de médecine et de psychologie : "Excerpta Medica, Index Medicus, IPA (International pharmaceutical abstracts)".

Je me suis servi des index matières avec les entrées :

**speech recognition
speech recognition process
speech recognition machines
signal treatment
speech synthesis**

Cette recherche n'apporta aucun résultat.

En revanche, si une recherche dans le bulletin signalétique Pascal fut également infructueuse (section "psychologie, psychopathologie, psychiatrie"), j'ai trouvé dans le périodique "Pascal explore" de nombreuses références recensées tous les mois sous l'entrée :

Reconnaissance et synthèse de la parole et du son.

Mais, plus tard, la délimitation exacte du sujet m'a convaincu de leur inutilité pour mon travail, puisqu'elles se rapportent toutes au côté technique, informatique de la question.

c) Bibliothèques interuniversitaires, section lettres :

Menée à Lyon et Grenoble avec les mêmes entrées-matières qu'en section médecine, la recherche fut vaine.

d) Bibliothèques interuniversitaires, section sciences :

Quelques références ont été trouvées à Lyon et Grenoble avec l'entrée :

Traitement du signal

... mais elles étaient trop générales. Par contre une thèse de 1979 sur la reconnaissance automatique du locuteur fut précieuse pour la délimitation du sujet.

e) Institut de phonétique de Grenoble III :

Sur les conseils de M. Lavorel qui m'a mis en contact avec un enseignant de l'Institut de phonétique de Grenoble, J. L. Boe, j'ai obtenu de celui-ci une bibliographie qu'il a écrite avec D. Pascal et J. Raymond :

Les caractéristiques individuelles de la voix : bibliographie indexée.

L'étude de ce document, riche de 1261 références a provoqué la phase finale de délimitation du sujet évoquée ci-dessus.

En effet, l'indexation de ce document nous a permis de dégager des références concernant l'identification juridique du locuteur que j'ai croisées avec celles indexées à :

Vérification du locuteur
Age
Reconnaissance du locuteur
Sexe
Anatomie.

Cependant, la plupart des références obtenues provenaient de rapports de congrès d'origine anglo-saxonne. Il était donc nécessaire de les commander à la British Library de Londres. Si cette démarche est possible par le prêt inter-bibliothèques, il apparaît que les demandes d'actes de congrès sont assez longues. Aussi avant de les effectuer, j'ai essayé de localiser en France ces documents, absents des bibliothèques universitaires. Une interrogation du Catalogue collectif national, à propos de la localisation d'un périodique de médecine légale, m'a permis de connaître l'existence de la bibliothèque de l'Ecole Nationale Supérieure de Police, à Saint Cyr au Mont d'Or (Rhône).

f) Ecole nationale supérieure de police (ENSP) :

La bibliothèque de cet établissement n'est pas libre d'accès mais une demande écrite et motivée m'en a ouvert les portes. Si l'ENSP possède de nombreux ouvrages, leur approche n'est pas facile. En effet, la bibliothèque n'était installée dans de nouveaux locaux que depuis trois semaines et était en cours de réorganisation. La gentillesse de son personnel a pallié, en partie, ce désagrément.

Ma stratégie, tendant à éviter une demande de photocopies d'articles à la British Library, ne fut qu'à moitié justifiée. En effet, si je n'ai pas trouvé les actes des congrès que je cherchais, en revanche, j'ai obtenu une partie des articles qui m'intéressaient traduits dans des revues professionnelles de police.

Quatre périodiques ont spécialement retenu mon attention :

Revue internationale de police criminelle, publication officielle d'Interpol.

Revue internationale de criminologie et de police technique.

Revue de science criminelle et de droit pénal comparé.

Revue de médecine légale et de toxicologie devenue en 1981 la revue de médecine légale et de droit médical.

J'ai procédé à un dépouillement systématique des sommaires, à partir de 1947 pour la plus ancienne des quatre, et me suis servi des index matières annuels, pour les trois qui en possédaient, en utilisant les entrées :

Documentoscopie

Expertises

Identification

Médecine légale

Police scientifique

Technique policière

A la suite de ce travail en bibliothèque, j'ai rencontré l'enseignant de l'école qui s'intéresse plus particulièrement aux techniques de police scientifique : le commissaire divisionnaire Schir. Un entretien avec lui m'a permis de constater que la rareté des références trouvées n'était pas due à un mauvais approvisionnement de l'école, mais à des lacunes dans ce secteur particulier de la recherche.

Sur la recommandation de M. Schir, j'ai contacté au Ministère de l'Intérieur (Direction de la police judiciaire, sous-direction de la police technique et scientifique) M. Barthe, dont le service s'intéresse particulièrement à l'identification de la parole. Il ne m'a pas apporté, pour l'instant, plus d'informations, la personne suivant ce dossier ayant quitté son poste depuis trois mois. J'attends de lui quelques compléments bibliographiques non parvenus à la date de la rédaction du présent projet.

g) La mise à disposition des documents :

J'ai eu recours au prêt entre bibliothèques pour disposer de monographies en langue anglaise et d'actes de congrès américains. L'identification exacte de ces derniers a été permise avec :

**Index of conference proceedings received 1976-1978.
The British Library. Lending Division.**

Comme un certain nombre de références étaient tirées des actes d'une conférence régulière se tenant à Lexington (Kentucky) sur les contre-mesures en matière de lutte contre le crime et qu'elles s'arrêtaient en 1977, j'ai voulu commander les sommaires des actes des conférences suivantes. Mais, la consultation des volumes annuels de "Index of conference proceedings received" m'a permis de me rendre compte que ces assemblées ne s'étaient plus tenues à compter de juillet 1977.

B) Recherche automatisée :

Traiter un sujet portant sur plusieurs disciplines est sans doute stimulant mais gêne certaines démarches. Voulant compléter ma recherche manuelle par une recherche automatisée, je me suis enquis des bases de données susceptibles de me convenir.

Les bases médicales ou scientifiques, telles Medline ou Pascal m'étaient fermées puisque je n'avais rien trouvé sur leur équivalent papier. J'ai donc consulté divers répertoires de bases de données :

Banques de données de l'université et du CNRS (1986, DBMIST).

Répertoire des banques de données en conversationnel (Association nationale de la recherche technique, 1987).

Annuaire international des banques de données juridiques, 1986-1987 (F.L.A., La Créatique).

Mais rien, dans les notices de ces ouvrages ne semblait se rapporter au sujet, ce que m'ont confirmé les conservateurs des bibliothèques interuniversitaires à qui j'ai demandé conseil.

3) Conclusion :

Si cette recherche débuta avec des certitudes, celles-ci se sont rapidement dissipées et le rassemblement des documents a été relativement difficile, étant donnée leur dispersion. Il a fallu éliminer beaucoup d'information non pertinente.

Cette sélection rigoureuse a abouti à une liste de 11 références. Seuls les travaux les plus récents de chaque auteur ont été retenus.

III Synthèse

Deux procédés dominent les recherches sur la reconnaissance de la parole :

- la méthode des sonagrammes.*
- l'identification de la voix par ordinateur.*

Pour mémoire, nous citerons aussi l'identification par l'audition de la voix, méthode subjective mais qui n'est pas sans intérêt. De nombreuses personnes sont capables de reconnaître des voix familières par leur timbre, leur élocution.

La relative fidélité de cette reconnaissance explique que, dans plusieurs affaires récentes, les autorités judiciaires l'aient permise, telle la diffusion le 17 mars 1983 à la radio de la voix des racketteurs d'André Schoch, un coiffeur ajaccien assassiné.

Cependant, la subjectivité de cette reconnaissance provenant du fait que les critères d'identification restent mal connus, que la voix peut être déguisée ou déformée dans la transmission, ne permet pas d'établir avec certitude une culpabilité.

1 - La méthode des sonagrammes :

A) Les sonagrammes du professeur Kersta :

Le professeur Lawrence G. Kersta est ingénieur aux "Bell Telephone Laboratories" (USA) :

a) Les recherches du professeur Kersta :

En 1963, au cours du Cycle international de Police scientifique qui se tint à Paris, sous les auspices d'Interpol, Kersta fait un exposé sur la possibilité d'enregistrer les caractéristiques de la voix sous la forme d'un spectrogramme (encore appelé selon les auteurs vocogramme ou sonagramme) en vue d'en faciliter l'identification.

Pour donner une approche de ce qu'est un sonagramme, rappelons brièvement une définition de la voix :

"La voix est constituée par l'ensemble des sons fournis par l'air sortant des poumons et qui traverse le larynx et les cavités buccales". L'homme ayant en plus la faculté d'articuler les sons (de les moduler en quelque sorte) possède ce que l'on appelle la parole.

Sous l'action de l'air poussé par les poumons, le larynx produit des sons articulés par la glotte et modifiés par les diverses cavités rencontrées sur le trajet (dont le "couplage" est modifié par la langue).

*La voix est constituée de sons variables dans le temps.
Un son est caractérisé par une fréquence fondamentale, celle-ci
étant accompagnée ou non de fréquences harmoniques.*

Le mécanisme de la voix

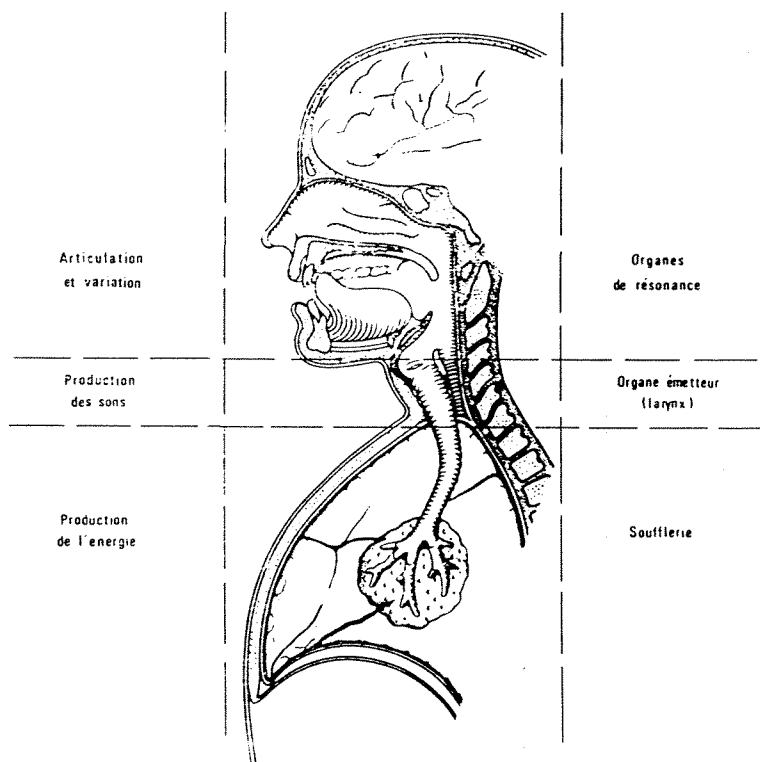
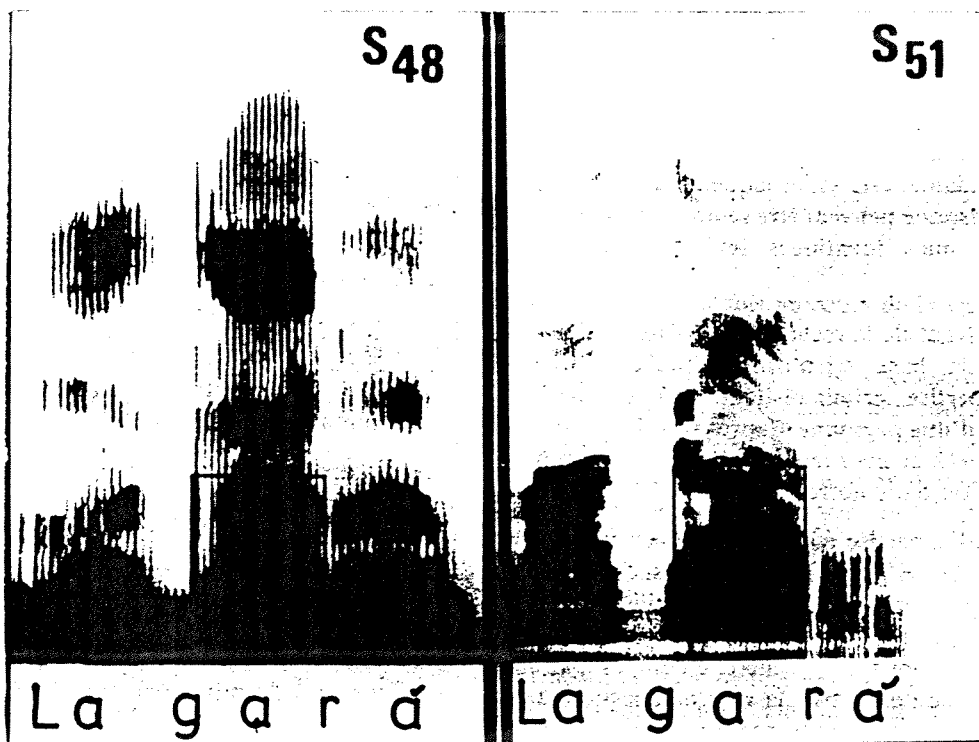


FIG. 1

On s'aperçoit ainsi qu'il est important d'examiner simultanément les trois paramètres (fréquence, intensité, temps) pour conserver la corrélation existante dans ce tryptique, et non d'isoler les paramètres entre eux si l'on souhaite obtenir de ce phénomène une représentation offrant le maximum de renseignements.

Le spectrographe (appareil produisant les spectrogrammes) permet une telle analyse. Les mots à étudier sont enregistrés sur une bande magnétique et, à l'aide de filtres, on balaye successivement le spectre de fréquence marquant à chaque instant l'énergie contenue (ou non) dans la largeur de bande considérée.

Le principe du sonagramme repose sur l'enregistrement sur papier du spectre de fréquence vocale. Sur le dessin obtenu, le temps figure en abscisse, les fréquences en ordonnée et le noircissement du papier est fonction de l'énergie recueillie.



Selon Kersta, les particularités d'une voix donnée permettent de la différencier de toutes les autres même si l'on compare les voix de membres d'une même famille, de jumeaux, d'un ventriloque... C'est un des principaux facteurs dont il faut tenir compte avant de construire un sténographe électronique, par exemple. Il n'y aurait donc pas deux spectrogrammes identiques.

Pour une personne déterminée, la configuration de la bouche, de la gorge et des fosses nasales fait que l'énergie vocale se concentre dans des bandes de fréquence bien définies, et tout déguisement de la voix est inutile ; les conditions restent également les mêmes si la personne s'exprime dans une langue étrangère. La perte des dents, l'ablation des amygdales, ou même le fait de "parler la bouche pleine" ne modifierait qu'un champ très réduit du spectrogramme de la voix.

Les expériences de Kersta ont consisté d'abord à "photographier" les voix de 15 hommes et de 10 femmes. On leur avait demandé de lire des phrases de 10 mots contenant des vocables tels que "me", "it", "you", d'usage fréquent au téléphone. Tous les sujets prononcèrent ces mots différemment. On établit ainsi les "vocogrammes" de ces mots-clés prononcés par les 25 personnes en question ; puis les images furent "battues" comme des cartes à jouer. Le spécialiste entraîna ensuite ses collaborateurs à les étudier et à les classer par individu et par mot-clé. Des 25 000 opérations effectuées, 97% furent correctes.

Kersta pense qu'il peut s'agir là d'un moyen de preuve presque aussi efficace que les empreintes, et qui permet de confondre des criminels beaucoup plus sûrement qu'en "auditionnant" directement les voix.

b) Critiques du travail du professeur Kersta :

α) Apports de J. P. Mamoux :

J. P. Mamoux, attaché aux services techniques de la préfecture de police de Paris, s'intéresse aux conclusions de Kersta qui affirme, nous l'avons vu, que chaque individu possède des caractéristiques vocales qui lui sont propres et qu'une déformation volontaire de la voix (déformation n'affectant pas physiquement et d'une manière permanente les organes phonatoires) n'altère pas les caractéristiques ou les points particuliers du sonagramme.

Pour vérifier cela, il a procédé à 13 enregistrements effectués par un imitateur connu et les a communiqués à Kersta. Celui-ci retrouvant quelques points caractéristiques dans ces 13 enregistrements, apparemment différents, a pu, avec une forte probabilité, affirmer que les paroles avaient été prononcées par la même personne.

Si cette expérience plaide plutôt en faveur de Kersta, le problème est que la méthode de reconnaissance, s'appuyant essentiellement sur la forme physique des cavités, il est bien évident qu'un individu utilisant des organes artificiels (faux palais, prothèses...) pour modifier d'une façon appréciable ses propres organes de phonation, produirait probablement un sonagramme fondamentalement différent de celui obtenu naturellement.

Pour Mamoux, cela infirme certaines assertions qui laisseraient supposer que le spectrogramme pourrait être utilisé en tant que preuve négative.

D'autre part, une autre restriction quant à l'évidence de la méthode, réside dans le fait que, dans certaines affaires policières où des enregistrements de conversations confidentielles se trouvent au dossier, les dits enregistrements sont souvent de mauvaise qualité.

En effet, dans la plupart des cas, ils ne sont pas effectués par des professionnels et, fréquemment, pour des raisons de discrétion, la prise de son est faite dans de très mauvaises conditions. Il en résulte que le rapport signal/bruit de la conversation restituée par le magnétophone est par trop faible, et que la reconnaissance des formes s'en trouve gênée par la présence du bruit qui se traduit sur le sonagramme par un fond grisâtre sur lequel ne se détachent pas les points singuliers recherchés.

La conclusion de Mamoux est que cette méthode d'identification de la voix n'est pas à rejeter a priori, mais de sérieuses études doivent être entreprises (dans toutes les langues) dans le but de définir certains critères et un système de comparaison automatique qui permettra, en appliquant la loi des grands nombres, d'indiquer de manière précise la valeur statistique qui s'attacherait à cette technique.

β) Remarques de P. Corsi :

L'idée d'empreinte vocale (voice print), qui pourrait jouer un rôle analogue aux empreintes digitales, a suscité de vives réactions que P. Corsi note dans sa thèse sur la reconnaissance automatique du locuteur.

1) *Ainsi Ladefoged et Vanderslice ont essayé de mettre sur pied une preuve expérimentale à l'encontre des thèses de Kersta, et lui reprochent d'avoir trompé "à la fois le public, la loi, et probablement lui-même", et rien de moins que d'avoir fait preuve de dogmatisme. A l'appui de cette réplique, sont les preuves expérimentales suivantes :*

- même en suivant la technique de Kersta, il est possible de produire des paires de spectrogrammes pratiquement identiques, provenant de prononciations de deux locuteurs distincts.

- même si les dires de Kersta étaient vrais, sa technique ne ferait que prouver ceci : les mots effectivement utilisés dans la procédure de comparaison ont été prononcés par le même locuteur ; mais justement, ces mots ont pu être copiés ou insérés dans la bande témoin. En outre, un locuteur peut, à différentes occasions, prononcer un mot de différentes manières, tandis que les variations d'un locuteur donné peuvent se superposer avec celles d'un autre. Selon les circonstances, deux locuteurs peuvent produire des sons dont les spectrogrammes sont très semblables.

Ladefoged et Vanderslice ont réalisé un contre-exemple de la manière suivante : 3 locuteurs hommes ont prononcé en chambre sourde des phrases durant environ 2 secondes. Ces auteurs ont trouvé qu'il était tout à fait possible de produire une "preuve" conduisant à une fausse identification de l'un des trois locuteurs, avec l'un des deux restants, ou aussi les deux à la fois, indifféremment.

2) Par ailleurs, Corsi signale qu'il a demandé l'avis de spécialistes de la parole aux Etats-Unis, et aussi d'autre pays. Parmi les nombreuses réponses, il n'y en eut pas une qui fut en faveur de l'utilisation de spectogrammes à des fins légales.

3) Enfin, une lettre de **Bolt, Cooper, David, Denes, Pickett et Stevens (1973)** expose l'état de l'art en reconnaissance à l'aide de spectogrammes et discute la possibilité d'application de ces méthodes. La conclusion réitère une mise en garde déjà formulée : la méthode n'a pas été clairement établie. Une réponse immédiate a été faite par **Black, Lashbrook, Nash, Oyer, Pedrey, Tosi et Truby (1973)**, qui met en cause la non-expérience personnelle des auteurs précédents, et affirme qu'un personnel entraîné est au contraire qualifié pour mener à bien les enquêtes.

B) Les sonagrammes du docteur Anghelescu :

a) Les recherches du Docteur Anghelescu :

Directeur de l'institut de criminalistique de l'Inspectorat général de la milice en Roumanie, Anghelescu utilise aussi les sonagrammes.

α) Remarques sur l'appareil phono-respiratoire :

Les recherches roumaines ont commencé par l'examen systématique des données de l'anatomie, de la physiologie et de la physiopathologie de l'appareil phono-respiratoire. Elles ont permis d'énoncer des hypothèses de travail quant à la possibilité d'établir le sexe d'une personne et d'identifier celle-ci d'après sa voix et son parler en roumain.

Les investigations ont démontré que l'hypothèse selon laquelle la voix et le parler permettent l'identification du sujet est scientifiquement fondée puisque :

- la voix reste stable au cours de la vie adulte jusqu'à la vieillesse.

- toutes les modifications vocales consécutives au vieillissement deviennent à leur tour des particularités concourant à l'identification.

- d'une personne à l'autre, il existe des différences de construction anatomique de l'appareil phono-respiratoire, du larynx, des cordes vocales, du pharynx et de la cavité bucco-nasales.

- la fonction phonatoire présente des particularités telles que la pression de l'air expiratoire sous-glottique, les différentes modalités de l'apparition et de disparition du son ainsi que la façon d'articuler.

Le timbre et la fréquence de la voix sont contrôlés par un automatisme acquis, et, par conséquent, durant l'acte phonateur, le sujet parlant effectue un contrôle seulement sur ce qu'il dit et non sur les particularités de sa voix. De même, des caractéristiques passagères ou définitives de la voix engendrées par certaines maladies de l'appareil phono-respiratoire deviennent des éléments d'identification pour la période postérieure à l'apparition de la maladie.

En résumé, on peut conclure, sur la base des recherches effectuées, que l'identification d'une personne d'après sa voix est possible grâce :

- à la stabilité de la voix pendant l'âge adulte, depuis la fin de la puberté jusqu'à la vieillesse, période pendant laquelle les caractéristiques générales et particulières dues à l'anatomie et à la physiologie de l'appareil phonatoire demeurent pratiquement constantes.

- aux différences de construction de l'appareil phonatoire d'un individu à l'autre.

- aux particularités de la fonction phonatoire.

Les recherches roumaines ont permis de formuler l'hypothèse fondamentale selon laquelle la voix et le parler d'une personne peuvent être semblables à beaucoup d'autres, mais identiques seulement à elles-mêmes.

β) La méthode d'expertise criminalistique :

L'objectif suivant de la recherche consista en l'élaboration, sur la base scientifique établie, d'une méthode d'expertise criminalistique qui permette l'identification d'une personne d'après sa voix, objectif atteint en 1972.

La méthode d'expertise établit avec exactitude les étapes du processus d'identification, leur succession, ainsi que les procédés et les techniques utilisés dans l'examen de la voix et du parler, et des modèles de voix soumis à la recherche comparative. Dans la comparaison des sonagrammes, il est tenu compte des éléments suivants : le vocabulaire, le dialecte, la manière propre de prononcer les mots de provenance étrangère, l'accent, l'intonation, les pauses, l'habitude de répéter certaines expressions à intervalles réguliers.

Quels en sont les résultats ?

Selon Angheliescu, les expertises effectuées par l'institut de criminalistique ont été confirmées dans la proportion de 99% par les tribunaux. Pour lui, le laboratoire d'expertise phono-criminalistique peut :

- établir l'authenticité d'un vocogramme de question.*
- déterminer le sexe d'une personne.*
- identifier une personne d'après la voix et le parler utilisés dans la conversation courante.*
- découvrir et démontrer la présence de déguisements ou d'imitations dans la voix et le parler.*

Lors d'une expertise, le laboratoire rend une des conclusions suivantes : "certainement positive", "certainement négative", "probable" et "aucune conclusion" lorsque les échantillons de question ne présentent pas un minimum de caractéristiques.

Les apports de l'ordinateur :

Comme les opérations à effectuer lors d'une expertise demandaient trop de temps, il s'est avéré nécessaire de réaliser un programme complexe qui permette de faire effectuer rapidement toutes les opérations par un ordinateur, et surtout d'obtenir une fonction d'estimation des caractéristiques phono-acoustiques établies par l'expert sur les sonagrammes, fonction qui puisse contribuer notablement à une plus grande exactitude des conclusions formulées.

Des résultats obtenus tant à partir de cas expérimentaux que de cas concrets, on peut conclure que ce programme permet une estimation très exacte des paramètres phono-acoustiques qui entrent pour beaucoup dans la formulation des conclusions d'identité, de probabilité ou de non-identité. L'ordinateur apporte donc une contribution particulière pour ce qui est de rendre plus objective la conviction de l'expert, formée lors de ses recherches en laboratoire, et la conclusion qu'il formule.

Les apports du laser :

Enfin, on en est venu à examiner les diagrammes en les éclairant à l'aide d'une source laser dans un système de filtrage optique, ce qui permet d'obtenir un spectre de Fourier (ce spectre est un diagramme indiquant la pression sonore en fonction du temps) qui synthétise toutes les caractéristiques de la voix transcrites sur les sonagrammes et qui offre la possibilité d'un examen comparatif intégral de ces derniers. Le pouvoir de synthèse de cette méthode permet d'obtenir des résultats exceptionnels, car le spectre offre une présentation schématique de toutes les caractéristiques phono-acoustiques présentes sur le sonagramme de la voix.

La méthode du filtrage optique au moyen du laser rend possible pour la première fois, du point de vue quantitatif et qualitatif, l'examen synthétique de toutes les caractéristiques de la voix. Celles-ci pourront être comparées par superposition, grâce au spectre de Fourier, de la voix de question (celle qu'on expertise) et de celle de comparaison (celle dont on connaît déjà l'origine), ce qui aboutira à la démonstration complète de leur coïncidence ou de leur divergence.

B) Critiques des travaux du docteur Anghelescu :

Ici encore, la nécessité d'une similitude des conditions de production et d'enregistrement de la voix du sujet et de celle supposée être la sienne fait que la méthode n'est pas applicable avec sûreté dans les affaires criminelles.

2 - L'identification de la voix par ordinateur : les travaux du Docteur Bunge

Le docteur Bunge est directeur du groupe de recherche scientifique au Bundeskriminalamt, à Wiesbaden (République fédérale allemande).

Si l'identification de la voix en utilisant la capacité de stockage de ses caractéristiques et leur comparaison au moyen d'un ordinateur ont déjà prouvé leur efficacité dans plus de 99% des cas quand le sujet se prête à l'expérience et que le signal sonore n'est pas déformé par la transmission téléphonique, par contre les études récentes comme celle de E. Bunge cherchent à élaborer des systèmes tenant compte de la mauvaise qualité des enregistrements recueillis dans les affaires judiciaires où les voix sont souvent déguisées et déformées par la transmission.

Le système AUROS (automatic recognition of speaker) parvient à un taux d'erreur de 1% des procédures de complexité moyenne dans des conditions standards quand la coopération du sujet est assurée.

Lors de l'identification des voix par ordinateur, deux problèmes fondamentaux se présentent :

- la non reproductibilité de la parole humaine qui fait que personne ne peut prononcer une phrase de manière identique (voir schéma).

- la quantité de données à traiter dans de courts intervalles de temps.

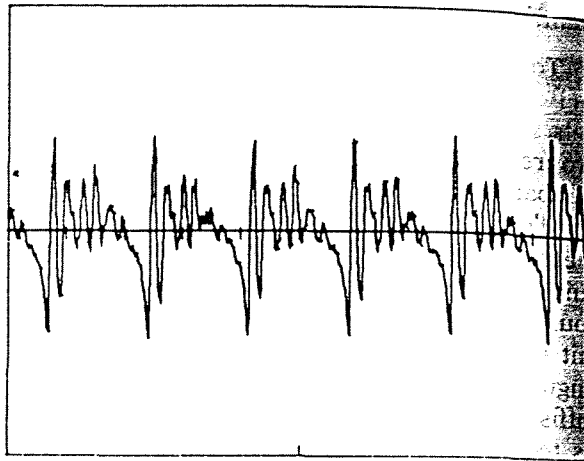


Figure 1

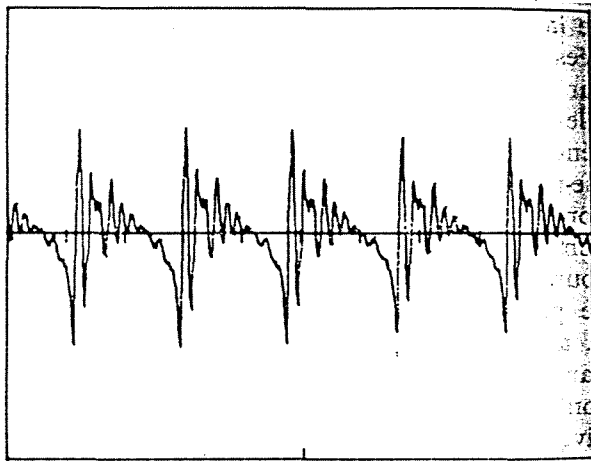


Figure 2

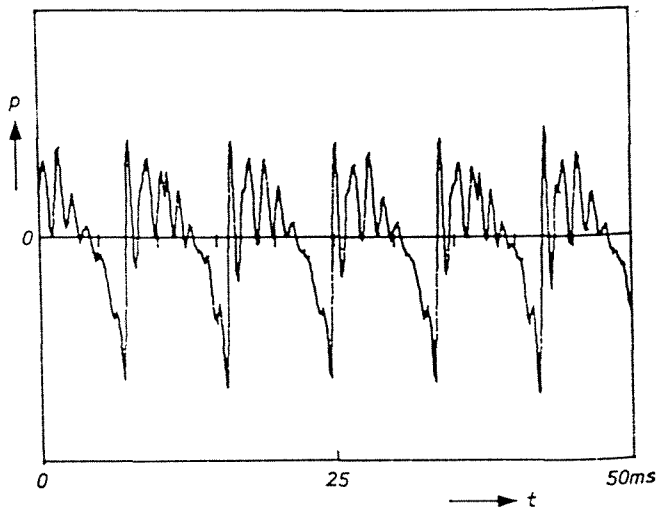


Figure 3

Pression sonore p en fonction du temps t pour trois émissions de la voyelle "a" par la même personne. Les différentes formes de la courbe montrent que la parole humaine n'est pas exactement reproductible ; cela pose un problème pour l'identification informatisée de la voix.

C'est pour cela que la voix d'un sujet ne peut être décrite par une seule émission. Il est nécessaire de procéder à une série d'émissions pour déterminer statistiquement les traits caractéristiques d'une voix. On utilise pour cela les techniques de reconnaissance des formes. Une seconde de discours humain peut être représentée par une séquence de 10 000 échantillons, et c'est l'ensemble de cette information traitée par l'ordinateur qui permettra de retenir les éléments caractéristiques.

La procédure suivie avec le système AUROS comprend trois phases :

1 - Le prétraitement du signal sonore de la voix associe plusieurs techniques : l'analyse d'un court segment présélectionné de l'émission de la voix, la mesure des tracés tels que ceux de la hauteur du son ou l'énergie déployée pendant une émission spécifique, enfin le calcul de la moyenne des vecteurs typiques de segments temporels successifs.

2 - La mise en place du système nécessite la combinaison des résultats d'analyse d'un certain nombre d'émissions vocales (au moins deux) et la comparaison à une voix de référence typique.

3 - Enfin le contrôle de fiabilité s'établit en reprenant les éléments non utilisés dans les phases précédentes et en les classant.

Ces recherches en cours sont d'un grand intérêt déjà pour une utilisation commerciale et industrielle de la reconnaissance de la voix par des machines.

Ceci s'effectue avec la participation volontaire du locuteur et des conditions standards d'émission et de réception de la voix. Cependant les variations des paramètres de la voix et de sa transmission dans les affaires judiciaires liées surtout à la non coopération du locuteur, les altérations de la transmission et les dégradations du signal sonore ne sont pas encore suffisamment connues. Des recherches fondées sur des études statistiques expérimentales sont menées notamment par l'équipe de criminologie de Wiesbaden.

3 - CONCLUSION :

Aujourd'hui, l'identification des algorithmes caractéristiques de la voix du locuteur et leur mesure est possible. Cependant, les modifications de ces paramètres dans les conditions de médecine légale, qui sont en particulier l'absence de coopération du locuteur et les dégradations du message dans la transmission par téléphone ne sont pas suffisamment connues pour que la reconnaissance du locuteur puisse faire fonction de preuve légale.

IV Bibliographie

1. [s. a.] *Les spectrogrammes vocaux*. **Revue internationale de police criminelle**, juin juillet 1965, no 142, p. 179. roles en graphismes.

2. **ANGHELESCU, I.** *L'expertise criminalistique de la voix*. **Revue internationale de police criminelle**, août septembre 1985, no 390, p. 180-185.

3. **BUNGE, E.** *Automatic speaker recognition by computers*. **International conferences on crime countermeasures**. Col of engineering, Lexington, University of Kentucky, may 1975, p. 23-28.

4. **BUNGE, E.** *Identification judiciaire de la voix par ordinateur*. **Revue internationale de police criminelle**, novembre 1979, no 332, p. 254-290.

5. **CORSI, P.** *Reconnaissance automatique du locuteur : présentation générale, méthodologie et expérimentation, perspectives d'application*. [Grenoble] : 1979. 289 p.

Th. doct. ing. : génie informatique : Grenoble, INPG.

6. GAYDA, M. , VACOLA, G. *L'identification de la voix humaine*. Journal de médecine légale, droit médical, 1983, T. 26, no 6, p. 715-719.

N, H. *Acoustic analysis of voice disguise related to voice identification*. International conference on crime countermeasures. Col of engineering, Lexington, University of Kentucky, april 1977, p. 31-35.

8. MAMOUX, J. P. *Identification de la voix humaine*. Médecine légale et domaine corporel, 1971, no 4, p. 35-38.

9. MARIANI, J. *La reconnaissance de la parole*. La recherche, octobre 1985, no 170, p. 1214-1226.

10. REITBOECK, H. *Speaker identification over telephone transmission channels*. International conference on crime countermeasures. Col of engineering Lexington, University of Kentucky, 1977.

11. TOSI, O. *Voice identification : theory and legal applications*. Baltimore : University park press, 1977.



