

## SYNTHESE DE LA PAROLE

### COMPTE RENDU D'UNE SEANCE DE TRAVAIL. SUR L'ELECTRICAL VOCAL ANALOG ( EVA III )

#### ESQUISSE D'UNE METHODE

par GUTH G. et BURGSTÄHLER D.

« La Nature écrit Aristote, comme nous avons coutume de dire, ne fait rien en vain et l'homme est le seul être vivant qu'elle ait doté du venin qu'est la parole... ». La singularité de ce don jointe au caractère merveilleux, mystérieux et sacré du langage font que l'acte de la parole a été ressenti comme une énigme par les civilisations de tous les âges. Aussi a-t-on cherché à étudier la voix, sa formation et sa production dès les temps les plus anciens.

— Les philosophes grecs avaient établi déjà une classification très savante des sons de leur langue (1) et pour sauvegarder la prononciation correcte de leurs livres sacrés les Védas (2), les Hindous s'étaient appliqués de leur côté et dès la plus haute Antiquité, à déterminer la position et le fonctionnement des organes au cours de la phonation (3). Les Arabes ne furent d'ailleurs pas en reste : pour préserver la récitation du Coran ils se donnèrent également, à une date très ancienne une phonétique descriptive de leur langue. Certains de leurs théoriciens comme AL-KHALIL et son disciple SIBAWAYHI ont exprimé d'ailleurs des idées que la phonétique moderne devrait reprendre et confirmer (4).

— La tentative inverse, qui prolongeant la démarche analytique, a abouti à l'époque contemporaine et grâce au développement de la science phonétique depuis plus d'un siècle à la réalisation de ce rêve tant caressé par l'humanité qui consiste à faire parler une machine, à fabriquer de la parole artificielle au moyen d'un synthétiseur, ne date pas d'hier non plus

#### BREF APERÇU HISTORIQUE

##### *Les machineries des Anciens...*

Dans la nuit des temps, les Dieux partagent avec l'homme le rare privilège que la nature lui a accordé, aussi la voix fut-elle considérée tout au long des âges comme une preuve de l'origine divine de l'espèce humaine. Don divin, la parole fut attribuée à la statue de Memnon le fils du soleil, dans la mythologie de l'Egypte ancienne. Pour faire parler leurs

1) cf. GEORGES STRAKA : *Album phonétique* p. 152 - Les presses de l'Université Laval, 1965.

2) HOLGER PEDERSEN : *The Discovery of language* - Indiana University Press London, 1959.

3) cf. G. STRAKA : *Album phonétique* pl. 51

4) cf. par ex. leur conception de la syllabe.

Divinités, les païens usaient déjà de savants subterfuges ! Les oracles, les voix miraculeuses font en effet intimement partie des civilisations grecques et romaines. A Delphes le Dieu parlait par la bouche écumante et convulsive de la Pythie qui montait sur un trépied placé au-dessus d'une ouverture d'où sortaient des vapeurs méphiques. Ailleurs la voix était transmise jusqu'à la bouche de l'oracle au moyen de tubes soigneusement dissimulés.

### *Un héritage des temps mythiques : Les « têtes parlantes »*

Le Moyen-Age chrétien avec sa mystique et son merveilleux, l'âge Baroque avec son goût du surnaturel ne déploieront pas moins d'ingéniosité à donner de la voix à des statues ou à des figures de leur création. Les réalisations les plus intéressantes furent sans doute à cet égard les « têtes parlantes » que confectionna à une époque plus tardive l'Abbé Mical. Ces têtes dont Rivarol a fait mention dans un pamphlet anonyme, la *Biographie Universelle*, le *Journal de Paris* de juillet 1783 ne nous sont, hélas, jamais parvenues, (l'Abbé Mical dit-on les aurait brisées lui-même), mais l'Académie des Sciences possède dans ses archives un rapport en date du 3 septembre 1783 signé de Le Roy, de Milly, La Place, Périer, Vicq d'Azir et Lavoisier qui nous renseigne de manière très précise sur la constitution de ces mécaniques et leur fonctionnement (bibl. 15). On y lit que « l'examen des machines de ce genre est curieux, parce que faites pour imiter la nature dans la prononciation des sons, elles peuvent aussi jeter quelque jour sur le mécanisme de la voix ». Les deux têtes d'apparence vénérable qui devaient prononcer sous la forme d'un dialogue « le Roi a donné la Paix à l'Europe - la Paix fait le bonheur des peuples » en dirent assez, selon le témoignage d'un contemporain, pour qu'on ne pût se refuser à leur accorder le don de parole... Avec sa « chambre à vent » ses résonateurs, le cylindre à picots, le mécanisme de leviers, de soupapes et de roues la tentative des « têtes parlantes » de l'Abbé Mical nous apparaît aujourd'hui comme un des produits les plus typiques de l'ère des automates qui s'est développée dans les milieux savants à l'époque du célèbre ouvrage de *L'Homme Machine* de La Mettrie. Déjà marquée de la raison des Encyclopédistes, elle porte encore en elle tout l'héritage mythique des siècles passés. Avec le développement progressif des arts mécaniques et l'éveil au XVIII<sup>e</sup> siècle de la pensée scientifique le rêve de faire parler une machine allait enfin devenir réalité... Pour les esprits nourris du *Discours de La Méthode*, la Synthèse apparut désormais aussi comme la contrepartie indispensable à toute démarche analytique.

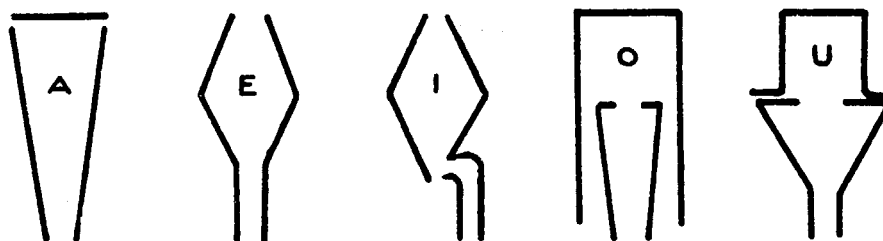
### AU SIECLE DES LUMIERES.

#### *Les résonateurs de Christian Gottlieb KRATZENSTEIN.*

En 1779 l'Académie des Sciences de Saint Petersburg offrit son prix annuel à qui préciserait les différences physiologiques existant entre les cinq voyelles [a], [e], [i], [o], [u] et construirait en même temps un instrument capable de les engendrer artificiellement. Ce fut le danois KRATZENSTEIN, professeur de physiologie à Copenhague qui obtint comme on sait le prix en présentant un mémoire sur la naissance et la formation des voyelles (5) et cinq résonateurs dont il avait étudié les profils et volumes, au moyen desquels il ren-

5) Ouv. publié en latin (*Tentamen Coronatum de Voce*) en 1780 à Pétersbourg et traduit intégralement dans les *Observations sur la Physique* (supplément de 1782, tome XXI p. 358-380).

força sélectivement certains harmoniques du son délivré par une anche et reproduisit les voyelles en question (fig. ci-dessous) :



Les resonateurs de KRATZENSTEIN (d'après LIENARD J.S. Bibl. 18)

Dans son mémoire, Kratzenstein donnait dès cette époque une description détaillée des organes de la parole et aboutissait à la conclusion que « les voyelles sont des sons modifiés par diverses ouvertures de la bouche et l'élévation de la langue » qui devait être déterminante pour les recherches ultérieures. Il ajoutait un peu plus loin : « Depuis quelques années, je me suis occupé dans mes moments de loisir, d'une machine qui pût contrefaire la voix humaine et qui comme un instrument de musique, pût par le secours des doigts, articuler des mots... ». La machine de KRATZENSTEIN, selon son propre aveu n'articulait encore que quelques syllabes mais il avait ouvert la voie à d'autres et en l'espace d'un demi-siècle des progrès considérables allaient être réalisés.

#### Le « parleur » de Wolfgang von KEMPELEN

Le baron Wolfgang von KEMPELEN qui avait déjà fait parler de lui pour avoir construit en 1773 une machine restée célèbre de nos jours, « l'automate joueur d'échecs », fut tenté à son tour « d'imiter par quelque instrument quelques sons de la voix humaine ». Stimulé par l'ouvrage publié un siècle plus tôt par le baron François Mercure Von Helmont, (6) l'idée qu'il y avait une liaison étroite entre la forme des lettres hébraïques et la position des différents organes phonatoires, qu'il ne manqua point de critiquer - von KEMPELEN s'instruisit des expériences de KRATZENSTEIN (qu'il cite d'ailleurs avec beaucoup d'admiration) et entreprit d'étudier à son tour le mécanisme de la parole et sa reproduction. Ses efforts aboutirent au bout d'une vingtaine d'années de patientes recherches à la création d'un véritable « instrument de parole », perfectionné au point qu'il pouvait lui faire dire « tous les mots latins, français et italiens qu'on lui proposait » (7), et à la publication à Vienne en 1791 d'un ouvrage relatif au Mécanisme de la parole, où il retraçait toute l'histoire de son entreprise, des premiers essais malheureux à l'évolution ultérieure de sa théorie (8) et qu'il accompagna aussi d'une description très détaillée de la machine de sa con-

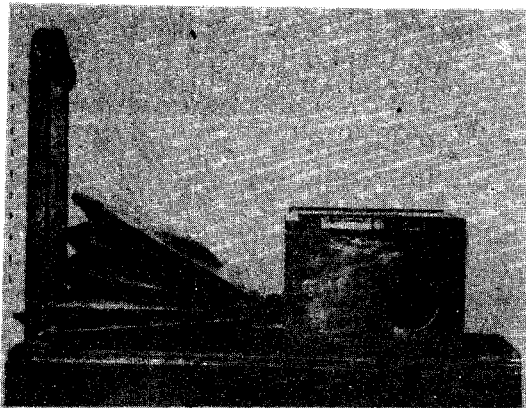
6) FRANÇOIS MERCURE VAN HELMONT : *Alphabeti Vere naturalis Hebraici Brevissima Delineatio* 1667.

7) cf. W. VON KEMPELEN : *Le mécanisme de la parole, suivi d'une description d'une machine parlante*, Vienne 1791 - « ... Toute imparfaite qu'elle soit elle donne au moins de bons principes pour en construire une plus parfaite. Enfin je l'ai portée au point que je lui fais prononcer d'abord et sans exception tous les mots latins, français et italiens que l'on me propose, les uns, il est vrai, mieux que les autres, mais du moins, plusieurs centaines de mots clairement et distinctement. Par exemple papa, maman, anatomie, Roma, maladie, sauté, astronomie, chapeau, Racine, soupé, charmante, opéra, comédie pantomime et aussi les mots longs et difficiles tels que Constantinopolis, Monomotapa, Mississipi, As-trakan, Anastasius, etc... »

8) cf. « Je dus donc rejeter entièrement un travail de près de deux ans et recommencer de nouveau. C'est ainsi que ma machine parlante et ma théorie de la parole ont fait des progrès égaux et que l'une a servi de guide à l'autre ».

ception. Son « parleur » que l'on peut voir encore à l'heure actuelle au Deutsches Museum de Munich, se compose comme cela apparaît sur la photographie et le schéma ci-dessous :

- d'un soufflet qui fournit l'air nécessaire à la production des sons (tout comme le font les poumons au cours de l'acte de la phonation),
- d'une anche vibrante capable de produire des oscillations de relaxation sous l'action du courant d'air, à la manière des vibrations glottiques.
- d'un entonnoir de gomme élastique représentant d'une manière bien imparfaite la bouche puisqu'il n'a point de dents, point de langue et point de partie molle du palais,
- d'une pièce de bois percée de deux trous munis de deux tuyaux de laiton symbolisant le nez,
- et d'une pièce centrale, le porte vent, qui se présente comme une boîte de 3 pouces  $\frac{1}{2}$  de long, 2 pouces  $\frac{1}{2}$  de large et 1 pouce  $\frac{1}{2}$  de haut reliant l'anche au soufflet et pouvant déboucher aussi sur l'extérieur grâce à un système de soupapes et de leviers par deux ouvertures dans lesquelles s'adaptent à gauche un tuyau coudé terminé par un entonnoir dont le couvercle ne laisse passer l'air que par une fente très fine pour la production du [S] et à droite un petit tuyau en bois en forme de bec de flûte qui constituait le sifflet du [ ]. Un troisième levier permet d'amener une aiguille de laiton au contact de la lame d'ivoire de l'anche qui la repousse par ses vibrations, subit des perturbations dans son fonctionnement et engendre ainsi un râclé qui donne à l'oreille l'impression d'un [R] « prononcé avec la partie molle du palais ».



Vue extérieure  
(Deutsches Museum Munich)

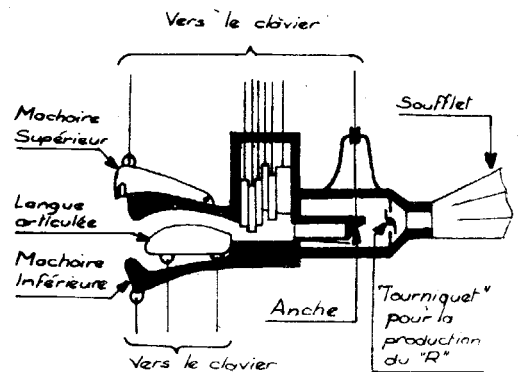


Schéma fonctionnel

*Le « parleur » de von KEMPELEN.*

Il fallait évidemment une assez grande dextérité pour faire parler la machine. Le bras droit devait actionner le soufflet. Les voyelles pouvaient être obtenues alors en faisant varier avec la main gauche dont deux doigts devaient s'appuyer sur les narines les caractéristiques du résonateur en gomme élastique placé en avant qui correspond à la cavité buccale. Pour le [i], la main devait être pressée contre tout le bord de l'entonnoir de manière à ne laisser qu'un petit passage à l'air ; pour la production du [a] elle devait être par contre entièrement détachée de l'ouverture. Par des positions intermédiaires de la main on pouvait obtenir de même toutes les autres voyelles ou « vocalités » car, à vrai dire, ces sons n'avaient qu'une certaine coloration vocalique, leurs timbres n'étant évoqués que par le déplacement d'un seul formant par rapport au fondamental (cf. Bib. 18).

En plus de la liquide [r], des sifflantes [s, z] et des chuintantes [ʃ, ʒ] pour lesquelles il suffisait d'actionner les leviers spécialement prévus à cet effet, la machine permettait aussi d'engendrer des sons à turbulence effacée comme [f] et [v] et la latérale [l]. [f] s'obtenait en fermant quasi totalement la bouche et en appuyant sur le soufflet de manière à faire fuser de l'air par les ouvertures et joints divers, [l] en entonçant le pouce jusqu'au fond de la bouche.

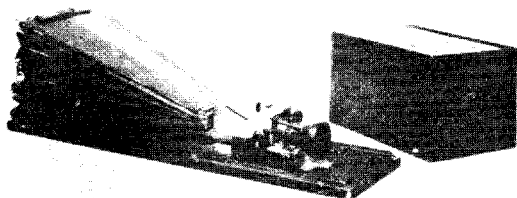
Les occlusives étaient produites d'une manière assez imparfaite, l'appareil ne permettant point de faire une distinction entre les bilabiales [p, b] les dentales [t, d] et les vélares [k, g]. Von KEMPELEN se contenta de former un son d'explosion unique en obturant complètement la bouche de manière à créer dans un premier temps une surpression que le soufflet additionnel avait pour rôle de renforcer et en retirant ensuite promptement la main de manière à libérer brusquement l'air emmagasinée. Il savait bien que la connaissance que nous avons du système d'une langue suppléerait aux insuffisances de la synthèse et puis il comptait aussi sur l'autosuggestion : « On est surtout aisément induit en erreur remarquait-il, quand on sait d'avance le mot que la machine doit dire et lorsqu'elle le prononce ajoutait-il, on s'imagine l'avoir entendu... ».

Tout comme les expériences de KRATZENSTEIN, ces premiers essais de synthèse de la parole de W. von KEMPELEN eurent le mérite, quelle que fût par ailleurs la qualité des résultats obtenus, de montrer que l'acte de la parole, aussi extraordinaire qu'il paraisse, se laissait réduire également à des mécanismes que la connaissance scientifique devait embrasser tôt ou tard. Aussi ces tentatives sont-elles à inscrire au nombre des œuvres de démythification entreprises par les philosophes et savants du XVIII<sup>e</sup> siècle, à la suite peut-être de ces découvertes récentes dont Fontenelle instruisait un siècle plus tôt, dans le cadre charmant d'un parc, la Marquise de la Méançère...

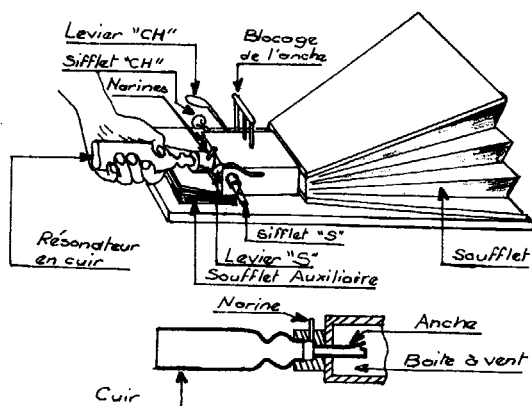
*Du XVIII<sup>e</sup> au XIX<sup>e</sup> siècle : de nouveaux constructeurs de synthétiseurs mécaniques.*

*La « speaking machine » de WHEATSTONE.*

Le « parleur » de W. von KEMPELEN fut pour l'époque un événement scientifique important et il bénéficia en ceci d'une publicité certaine. Son influence fut déterminante et son renom à l'étranger tel, qu'à Londres Sir Charles WHEATSTONE, professeur au King's College, construisait sur la base de ses observations une machine en tous points analogue, comme on peut en juger par la reproduction et le croquis ci-après :



Vue extérieure  
(Lent to Science Museum London)  
by King's College



Vue schématique  
(cf. bib. 5)

### « Speaking machine » de WHEATSTONE.

Cette machine a été conservée au Science Museum de Londres. WHEATSTONE fit aussi dès cette époque le point des connaissances relatives à la synthèse de la parole en publiant sur les machines parlantes un article que l'on peut lire dans *The Scientific papers of Sir Charles WHEATSTONE* où il rendait compte des travaux de von KEMPELEN, de KRATZENSTEIN, ainsi que des expériences que fit vers 1828 un autre physicien anglais, Robert WILLIS (cf. Bib. 25).

« Toute imparfaite qu'elle soit » avait écrit von KEMPELEN en parlant de sa machine, « elle donne au moins de bons principes pour en construire une nouvelle ». Il allait être à l'origine non seulement d'une imitation servile, mais du synthétiseur le plus perfectionné sans doute, que la mécanique ait jamais produit : « l'automate parlant » de Joseph FABER.

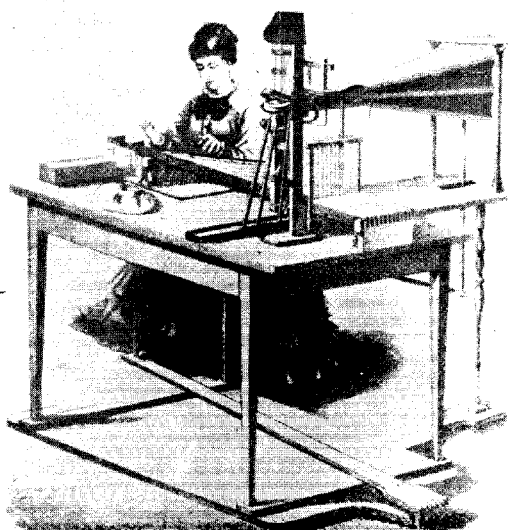
### L'automate parlant de Joseph FABER.

Joseph FABER, professeur de mathématiques à Vienne réalisa en effet vers 1835 une machine, qui s'inspirant dans ses grandes lignes du « parleur » de von KEMPELEN, présentait par rapport à celui-ci des perfectionnements et des avantages considérables :

- L'entonnoir en gomme élastique qui était sans doute la partie la plus faible de l'appareil de Von KEMPELEN fut entièrement abandonné et remplacé par un système plus adéquat pour simuler la bouche, un dispositif plus proche du modèle humain comprenant deux mâchoires mobiles en caoutchouc et une langue également mobile qu'un système de leviers permettait d'actionner.
- L'anche fut munie d'une rasette ce qui permit d'allonger ou de raccourcir la partie vibrante de la languette, de modifier la hauteur de la voix et de donner ainsi à la

machine de Von KEMPELEN qui disait tout sur un même ton « le véritable agrément » dont il rêvait (9).

Une série de diaphragmes coulissants permettaient en outre de donner à la cavité postérieure un profil particulier pour chaque voyelle et la machine avait sur le « parleur » de Von KEMPELEN l'énorme avantage d'être beaucoup plus maniable. Un clavier comportant quatorze touches reliées aux divers organes de l'appareil par un système de transmission particulièrement astucieux permettait en effet d'en user comme d'un harmonium. Gabriel, professeur à la Faculté de Médecine de Paris devait en donner d'ailleurs en 1879 une description très précise dans un numéro du Journal de Physique Théorique et Appliquée (cf. Bibl. 26).



Reproduction (d'après O. METTAS)

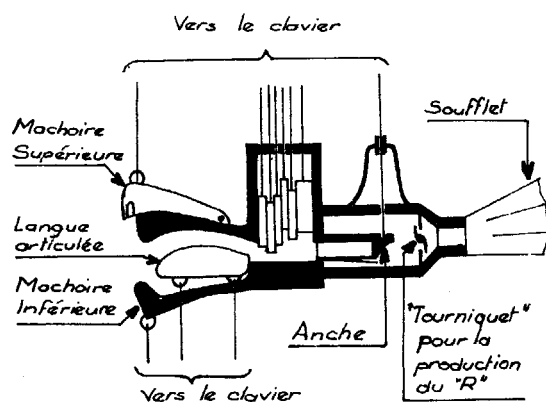


Schéma partiel (d'après DUMONCEL)

La machine fit avec son inventeur le tour de toutes les capitales d'Europe et devint très vite un sujet d'attraction. Quoi de plus étonnant d'ailleurs : non seulement elle pouvait produire des sons et des phrases intelligibles, parler à haute voix, mais encore chuchoter, et qui plus est, chanter « God save the Queen ! » L'enthousiasme général qu'elle suscita trouve un écho dans les journaux de l'époque : on y lit que « si elle n'a pas encore la voix d'une chanteuse d'opéra accomplie, elle laisse cependant loin derrière elle la machine de Von KEMPELEN. Ses sons mélodieux, si proches de la voix humaine, produisent une telle impression que l'on pourrait imaginer un frisson, une âme sensible se cachant à l'intérieur... ».

Devenu la propriété en 1850 du directeur des spectacles Barnum qui l'avait acheté pour son établissement, l'automate parlant et chantant de FABER fit alors un tour aux Etats Unis avant d'être ramené à Paris vers 1873.

9) cf. W. VON KEMPELEN : « ...obtenir sinon une sorte de chant, au moins une variation de la voix en parlant ce qui donnerait le véritable agrément à ma machine qui jusqu'à présent dit tout sur le même ton ».

*Evolution ultérieure des synthétiseurs mécaniques : de « l'automate » de FABER (1835) au « mechanical talker » de RIESZ (1937).*

Avec l'automate de FABER, l'année 1835 marqua une date dans l'histoire de la parole artificielle et des synthétiseurs. Les témoignages des contemporains prouvent assez à quel degré de perfectionnement on avait atteint aussi parassait-il désormais difficile de faire mieux avec les seules ressources des arts mécaniques. Plusieurs tentatives postérieures intéressantes et originales en leur genre méritent cependant d'être notées, en premier lieu celle d'Alexander Graham BELL.

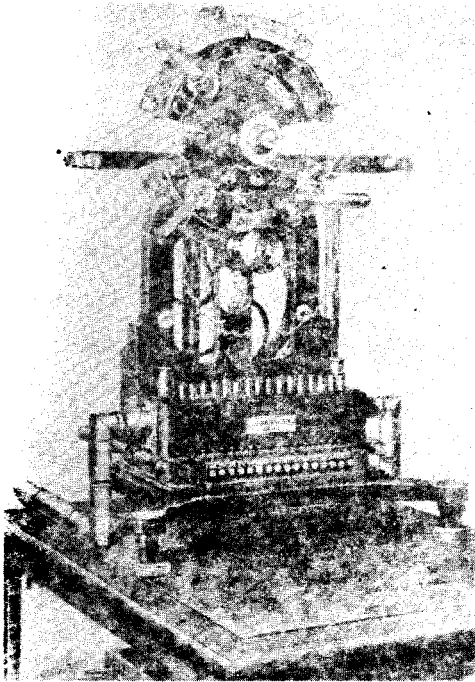
Au cours de son enfance passée à Edimbourg en Ecosse Alexander Graham BELL, l'ancêtre de la compagnie de téléphone de renom aujourd'hui mondial eut l'occasion de voir la reconstitution faite par WHEATSTONE du parleur de Von KEMPELEN. Il fut alors tellement impressionné qu'il décida de construire avec son frère Melville un automate parlant à lui. Conformément aux conseils que leur donna leur père Alexander Melville BELL, les deux garçons entreprirent de copier les organes de la voix en faisant un moulage en gutta serena d'après un crâne humain. Tous les organes furent représentés, les lèvres furent faites d'un support de fil de fer recouvert de coton et de gomme élastique ; les mâchoires furent réalisées en caoutchouc, et la langue simulée au moyen d'une pièce de bois rembourée de coton et recouverte de peau. Un soufflet fournissait l'air nécessaire à la production des sons, une membrane en caoutchouc munie d'une ouverture et tendue au-dessus d'un tube en métal tenait lieu de larynx et un système de transmission original permettait d'actionner l'ensemble. BELL put produire ainsi quelques phrases très simples. (cf. Bibl. 16).

A qui a le goût des anecdotes on raconte aussi très volontiers que durant son enfance, BELL, déjà fort intrigué par les mécanismes de la parole et sa production, expérimenta avec son « skye terrier ». Après avoir commandé au chien de s'asseoir et de grogner, il manipulait avec ses mains les diverses parties du conduit vocal de la bête de manière à faire varier les caractéristiques des résonateurs... et à produire de la parole... A son répertoire le chien avait les sons [a], [u], et [ou], les syllabes [ma] et [ga]. Selon la traduction orale, sa plus grande performance linguistique aurait été en outre « How are you Grandmamma ?.. » (cf. Bibl. 16).

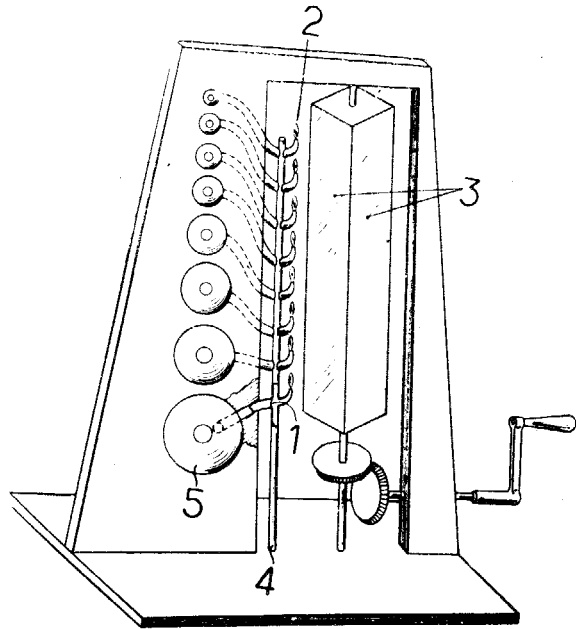
La seconde partie du XVIII<sup>e</sup> siècle et le début du XX<sup>e</sup> furent sans doute très fertiles encore pour le sujet qui nous préoccupe, *mais les essais de parole artificielle* devaient régresser pendant un quart de siècle environ au profit de la seule *synthèse des voyelles*. L'acte de la parole avait été démystifié et il importait moins désormais de faire parler une machine que de parvenir comme dans les autres disciplines scientifiques à une synthèse satisfaisante de l'objet d'étude considéré. Aussi peut-on voir dans cette évolution un aboutissement de l'esprit cartésien.

Des sons vocaliques synthétiques furent obtenus en 1875 par HELMHOLTZ au moyen d'une série de *résonateurs* entretenus par *des diapasons* et par KOENIG en 1882 au moyen d'une *sirène à ondes* de sa conception (fig. ci-après).





La sirène à ondes de KOENIG (1882)



Appareil de KOENIG pour l'analyse du timbre par flammes manométriques.

Ces appareils mis au point par des physiciens, des acousticiens, devaient intéresser au plus haut degré les pionniers de la phonétique instrumentale alors naissante (10). L'Abbé Rousselot consacra en effet à la description de ces machines et à leur fonctionnement une section assez importante des *Principes de Phonétique Expérimentale* (11) et son style reflète très souvent l'enthousiasme qu'il devait éprouver devant ces inventions. Il est à remarquer aussi au sujet de la démarche à la fois analytique et synthétique qui est propre à l'esprit mathématique et scientifique que HELMHOLTZ s'était servi au préalable de ses résonateurs pour l'analyse des sons et que KOENIG avait réalisé avant la sirène à ondes un appareil pour l'analyse des timbres par les flammes manométriques.

De fait l'analyse et la synthèse allaient progresser désormais à pas égaux. Bien d'autres essais de reconstitution des voyelles seraient encore à commenter, ceux qu'entreprirent par exemple MILLER (12) et STUMPF (13) avec des séries de tuyaux d'orgue, PREECE et STROH (14) au moyen d'un train d'engrenages réalisant un enregistrement phonographique, Sir Richard PAGET (15), auteur d'un grand nombre d'appareils, au moyen de résonateurs toujours visibles d'ailleurs à la Royal Institution de Londres, mais la place nous manque

10) La fondation au Collège de France de la première Chaire Phonétique date de 1897 - Elle fut occupée par l'Abbé Rousselot.

11) L'Abbé Rousselot; *Principes de Phonétique Expérimentale* H. Didier - Paris, 1924.

12) D. C. MILLER; *Sciences of Musical Sounds* (1916)

13) C. STUMPF; *Die Sprachlaute* (1926).

28, 1879.

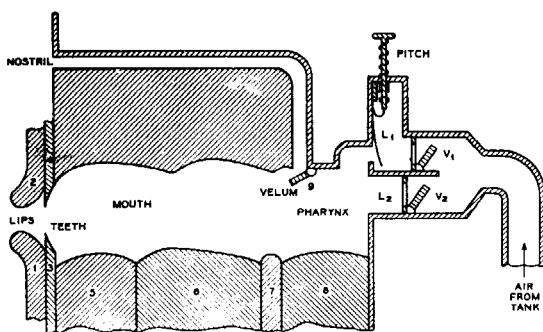
14) W. H. PREECE; A. STROH; *On the synthetic examination of vowel sounds Proc. Royal. Soc. London*

15) PAGET; SIR RICHARD; *Human Speech* - London and New York 1930.

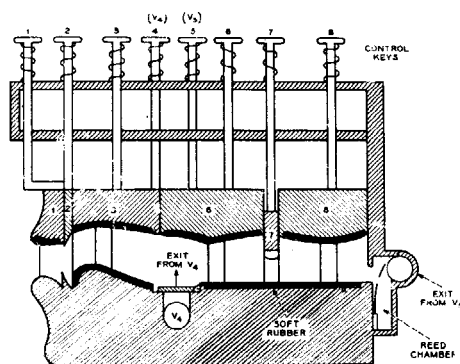
ici. Un dernier appareil mérite cependant d'être évoqué, le « mechanical talker » de RIESZ. Avec lui se termine l'ère des synthétiseurs mécaniques...

*La fin d'une ère : le « mechanical talker » de RIESZ.*

Après un quart de siècle d'interruption, d'essai de synthèse des sons isolés, de maturation aussi, on en revint dans la première moitié de ce siècle au continuum sonore de la chaîne parlée, à la parole artificielle. Sur le modèle des automates du XIX<sup>e</sup> siècle, RIESZ réalisa vers 1937 un appareil qui peut être considéré comme le dernier des synthétiseurs mécaniques.



*Schéma d'ensemble*



*Vue partielle*

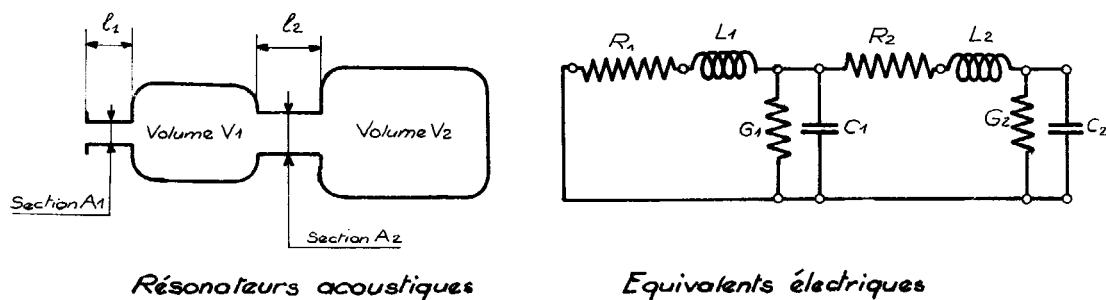
Le « mechanical talker » de RIESZ (1937)  
(d'après J. L. FLANAGAN)

La machine se présente comme une tentative de reconstitution intégrale, au moyen de matériaux et de dispositifs divers, du tractus vocal et de l'appareil phonatoire dans son ensemble. Tous les organes y sont représentés : les cavités pharyngale et buccale, le voile qui se trouve à l'entrée d'un conduit de dérivation nasal, la langue, les lèvres et même les dents. Une série de huit pistons munis de ressorts permettent de modifier à volonté la configuration du conduit, de faire varier les caractéristiques des résonateurs, de « jouer » ainsi avec cet instrument et de le faire parler... L'air nécessaire à la production des sons est fourni sous pression par une bouteille. Deux valves V1 et V2 permettent de contrôler l'admission : lorsqu'on ouvre V1 le courant d'air fait vibrer la lame L1 qui tient lieu de larynx et on peut produire alors tous les phonèmes sonores ; lorsqu'on ouvre V2 l'air excite directement le tractus vocal tout comme dans la production des sourdes lorsque les cordes vocales sont écartées. Il est possible en outre de modifier la longueur de la partie vibrante de la lame et de faire varier par conséquent la fréquence du fondamental. Maniée par un opérateur habile, cette machine était capable de produire selon un témoin une parole cou-

rante et intelligible. Le mot « cigarette », en particulier, pouvait être considéré comme une réussite. Le « mechanical talker » de RIESZ devait être le dernier des synthétiseurs mécaniques. L'électricité, depuis longtemps, avait fait son apparition et avec le développement de l'électronique qui offrait aux constructeurs des possibilités sans cesse croissantes, s'ouvrait pour la parole artificielle une ère nouvelle...

#### *La synthèse de la parole à l'ère de l'électronique.*

Les premiers essais de simulation électrique de la phonation datent des années vingt. Partant du principe qu'on peut associer à tout résonateur acoustique un circuit électrique équivalent (fig. ci-après) :



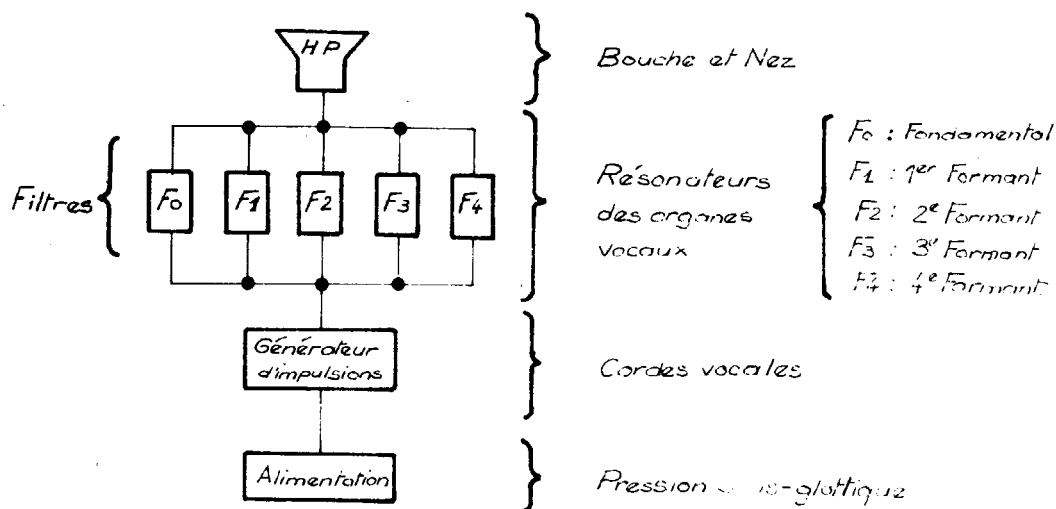
John Q. STEWART réalisa en 1922 le premier synthétiseur analogue électrique des organes vocaux. Les cavités buccales et pharyngales furent représentées par deux circuits électriques auxquels on appliqua un courant intermittent établi par un vibreur tenant lieu de glotte et les oscillations électriques furent transformées en vibrations sonores par un récepteur de téléphone.

En faisant varier la fréquence d'interruption du courant on pouvait modifier la hauteur du son et en modifiant les caractéristiques des circuits équivalents on agissait sur le timbre du son produit. L'appareil n'était pas encore suffisamment perfectionné pour engendrer véritablement de la parole artificielle mais il permettait d'obtenir par synthèse et de manière satisfaisante tous les sons vocaliques.

Les progrès allaient être rapides. Harvey FLETCHER des Laboratoires de la BELL téléphone présenta dès 1924 un appareil de ce type à la New York Electrical Society. Il ne produisait encore qu'un nombre restreint de sons, mais savait déjà dire « Mama » et « Papa ».

K. W. WAGNER réalisa en 1936 un synthétiseur de voyelles (16). Le schéma de l'appareil est donné avec les équivalents physiologiques de ses composantes par la figure ci-dessous :

16) K. W. WAGNER : Ein neues elektrisches Sprechgerät zur Nachbildung der menschlichen Vokale, 1936.



*Synthétiseur de voyelles de K. W. WAGNER*

Un générateur d'impulsions produit à la manière du générateur vocal un spectre très complet.

Une série de quatre filtres sélectifs et réglables, reliés à des amplificateurs et correspondant aux résonateurs de l'appareil phonatoire permettent d'isoler et d'amplifier à volonté certaines fréquences composantes de ce spectre. Les signaux recueillis à la sortie sont envoyés sur un haut-parleur représentant la bouche et le nez qui transforme les oscillations électriques en vibrations sonores. L'appareil produit par synthèse tous les timbres vocaliques.

L'électronique appliquée à la production des sons du langage venait de faire ses preuves. On avait dépassé désormais le stade des essais. Le succès était imminent.

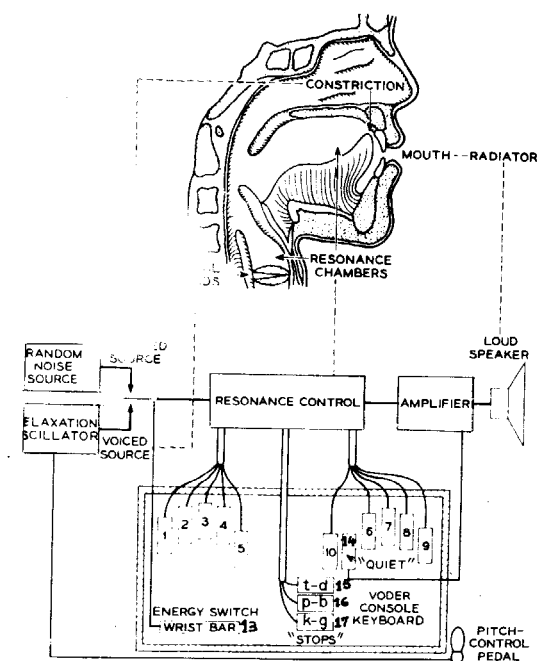
1939 - Aux Expositions mondiales de New York et de San Francisco.

Une grande première : *Le Voice Demonstrator (VODER)*.

L'année 1939 correspond pour la synthèse de la parole à une grande première. Aux Expositions Mondiales de New-York et de San-Francisco DUDLEY de la BELL Téléphone, RIESZ et WATKINS présentèrent une machine (17) sur les touches de laquelle une main féminine sténophonait de la parole : le Voice Demonstrator ou VODER dont le schéma et ses correspondants physiologiques sont donnés par la figure ci-dessous :

17) H. DUDLEY, R. R. RIESZ, S. WATKINS : *A synthetic speaker*, Journal of the Franklin Institute 227 - 1939.

## Schéma du Voice Demonstrator de DUDLEY, RIESZ et WATKINS

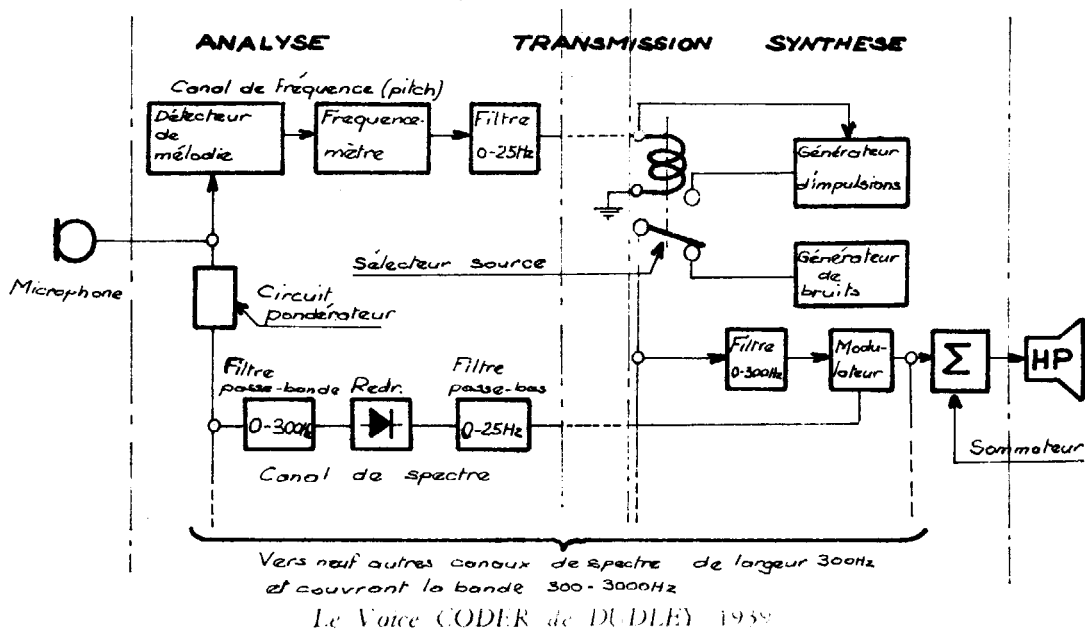


L'énergie nécessaire à la production des sons peut être choisie par la touche 13, commandée par le bras, soit sous la forme de bruit blanc, soit sous la forme d'impulsions périodiques selon que le conduit vocal est excité par le souffle comme c'est le cas dans la production des consonnes sourdes et de la voix chuchotée, ou par le son glottique comme c'est le cas pour les phonèmes sonores, les deux systèmes pouvant d'ailleurs fonctionner ensemble pour la production des consonnes sonores. Le signal issu du générateur de bruit blanc, de l'oscillateur à relaxation ou des deux sources à la fois, attaque ensuite un ou plusieurs des dix filtres passe-bande connectés en parallèle qui couvrent la gamme de 0 à 7500 Hz, et qui correspondent aux résonateurs du conduit vocal. La mise en circuit de ces filtres est commandée par les touches 1 à 10 qu'actionnent les doigts de l'opératrice. L'appareil comporte en outre une touche qui permet de baisser le niveau sonore de 20 db (touche n° 14) et trois autres (touches 15-17) permettant d'obtenir de brusques variations de niveau et de produire ainsi les occlusives alvéodentales (touche 15), bilabiales (touche 16) et vélares (touche 17). La hauteur du son peut être modifiée à volonté et contrôlée grâce à une pédale. « Converser » avec cet appareil n'était évidemment pas chose facile : l'opérateur devait être un très bon pianiste, avoir une très bonne connaissance de la structure phonétique de la langue utilisée et beaucoup d'entraînement car la continuité nécessaire à la compréhension du message n'était maintenue que par sa dextérité et son habileté. De ces difficultés est née sans doute l'idée d'une commande automatique du VODER qui a abouti à la construction du VOCODER.

*Premier d'une longue série : le Voice CODER de DUDLEY .*

En même temps qu'il présentait le VOCODER, DUDLEY mettait au point un système de reproduction automatique de la parole : Le Voice CODER. L'appareil se présente comme un VODER doublé d'un analyseur : au lieu d'être actionné par la main d'une opéra-

trice, il est commandé par les signaux provenant de l'analyse de la voix à transmettre. Son schéma de principe est donné par la figure ci-dessous :



La parole transformée par un microphone en un signal électrique est appliquée à l'entrée de l'analyseur. Celui-ci se compose d'une série de dix filtres branchés en parallèle qui couvrent une bande de 0 à 3 000 Hz et qui sont suivis chacun d'un redresseur et d'un filtre passe-bas avec une  $F_{lim.} = 25$  Hz. De cette manière, les 10 canaux transmettent non plus les vibrations sélectionnées par chacun des dix filtres, mais seulement les signaux correspondant à 25 variations d'amplitude de ces vibrations par seconde, autrement dit, les variations lentes de l'enveloppe oscillographique (Bibl. 10).

La parole est ainsi « codée » conformément à la dénomination de l'appareil et le « codage » constitué de l'ensemble de ces variations sert : de « programme » à la synthèse.

Les signaux issus de ces dix canaux commandent en effet un nombre identique de modulateurs fonctionnant comme des « portes » électroniques pour la série des dix filtres similaires dont se compose le synthétiseur ; le débit de chacun est fonction de la sorte de l'importance qu'a à l'analyse chaque bande de fréquence considérée. Ces filtres de synthèse peuvent être attaqués soit par un oscillateur à relaxation, (production des sons vocaliques et des consonnes sonores), soit par un générateur de bruit blanc (production des sons consonantiques). Le choix de l'une ou l'autre de ces sources est automatique : il relève du fréquence-mètre de l'analyseur qui agit par l'intermédiaire d'un relais et qui a la double fonction d'établir une distinction entre les sons et les bruits et d'enregistrer les variations de la fréquence fondamentale de la voix. Les tensions partielles obtenues à la sortie des canaux du synthétiseur s'additionnent et attaquent un haut-parleur à travers un amplificateur commun. La parole est ainsi reconstituée.

Cet appareil dont la bande passante totale était réduite en définitive à moins de 300 Hz permettait d'obtenir encore une parole intelligible à 90 % ! Ce fait devait intéresser au plus haut degré la technique des télécommunications. Il apparaissait en effet, que la bande passante du téléphone (300-3 000 Hz) était très largement excédentaire pour transmettre le contenu sémantique du langage et que l'emploi d'un appareil de ce type permettrait théo-

riquement de multiplier par dix le nombre des communications, rien qu'en utilisant les lignes existantes. (Bibl. 18). Avec le VOCODER la synthèse de la parole déboucha ainsi sur une application pratique de première importance.

En raison même de l'intérêt qu'il présentait, l'appareil de DUDLEY connu depuis son invention d'innombrables perfectionnements. Construit actuellement par la BELL téléphone Inc. (N. Y.), le VOCODER a donné lieu aussi à de très nombreuses variantes si bien que l'on applique aujourd'hui ce terme générique à plus d'une dizaine d'appareils : Multiplexing Channel Vocoder (Vilbig et Haase 1956) Pattern-matching Vocoder (Dudley 1958), Auto-correlation Vocoder (Schroeder 1959-1962) Formant Vocoder (Stead et Jones) etc... (Bibl. 5)

*Synthèse de la parole par « relecture » des sonagrammes à l'aide d'appareils « play-back »*

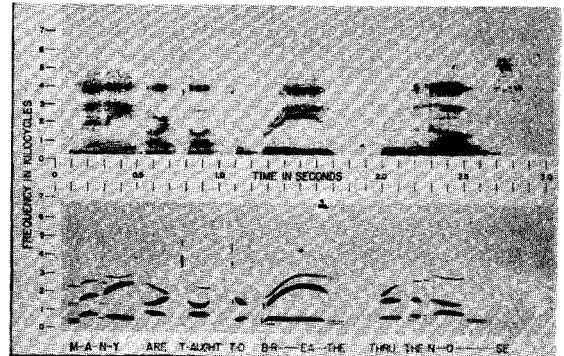
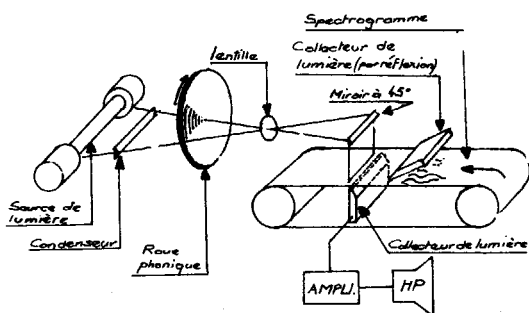
Le « PATTERN PLAYBACK » de F. S. COOPER.

Le développement progressif des méthodes d'analyse et la mise au point dans la période de l'après-guerre de cet appareil tant attendu de tous ceux qui s'étaient livrés jusqu'alors à l'étude mathématique des courbes oscillographiques, que fut le *SonaGraph* donnèrent aux recherches phonétiques et à la synthèse de la parole en particulier une impulsion nouvelle.

Le *Sound-spectrograph*, comme on l'appelait encore, fut décrit par W. KOENIG, H. K. DUNN et L. Y. LACY de la BELL téléphone dans *The Journal of the Acoustical Society of America* en juillet 1946 (vol. 17). Dans un ouvrage resté célèbre, intitulé « *Visible Speech* » R. K. POTTER ; G. A. KOPP et H. C. GREEN publièrent dès l'année suivante les premiers résultats spectrographiques des sons de l'anglais. L'analyse qui avait connu depuis l'époque de KOENIG (le constructeur de la sirène à onde et de l'appareil à flammes manométriques) un développement plus lent que la synthèse offrait désormais à la parole artificielle des données beaucoup plus complètes, des informations beaucoup plus sûres, des possibilités tout à fait nouvelles et même des enregistrements directement utilisables pour recréer la voix ! Le sonagramme étant en effet une représentation adéquate du phénomène sonore, il devenait possible de reconstruire à partir de cette représentation tous les caractères de la parole au moyen d'un appareil capable de la relire, c'est-à-dire d'un «PLAY BACK»

Le premier instrument de ce type fut mis au point vers 1950 par les laboratoires HASKINS : c'est le « PATTERN PLAYBACK » décrit par F. S. COOPER (18) qui permet à partir de spectrogrammes peints à la main de reconstituer la parole conformément au schéma ci-dessous :

*Schéma de principe du PATTERN PLAYBACK des Laboratoires HASKINS*  
(d'après F. S. COOPER)



sonagramme original et modèle simplifié pour servir de programme à la synthèse (d'après J. L. FLANAGAN)

18) cf. FRANKLIN S. COOPER : Speech synthesizers in *Proceedings of the fourth International Congress of Phonetic Sciences Helsinki, 1961*, p. 7.

L'appareil se compose d'un générateur à fréquences multiples dont la pièce maîtresse est un disque transparent entraîné par un moteur et sur lequel sont disposées 50 pistes photographiques concentriques semblables à celles qui se trouvent sur le bord des films cinématographiques. Les fréquences enregistrées sur ces bandes sont choisies de telle manière que si l'on tourne le disque à vitesse constante, elles sont en rapport harmonique avec le fondamental fixé à 120 Hz. Le disque est éclairé par un tube et la lumière focalisée par une lentille cylindrique traverse ces bandes et se trouve divisée ainsi en 50 « pinceaux ». Ces faisceaux lumineux dont le premier est modulé à 120 Hz, le second à 240 Hz, le dernier à 6 000 Hz tombent à travers une lentille sur un miroir incliné à 45° qui les renvoie vers le bas sur une bande transparente sur laquelle ont été dessinées au préalable, à la gouache épaisse, les structures mises en évidence par l'analyse. Les 50 faisceaux qui font une largeur de 5 pouces correspondent de la sorte aux fréquences des 50 harmoniques d'un spectrogramme naturel de mêmes dimensions.

La préparation du spectrogramme artificiel a été décrite en détail par DELATTRE (Bibl. 6) : on peint pour chaque harmonique une ligne dans le sens du temps. La fréquence est d'autant plus haute que la ligne est plus élevée, la durée d'autant plus grande que la ligne est plus longue (7,2 pouces/s) et l'intensité d'autant plus forte que la ligne est plus large ou plus réfléchissante (largeur max. 1/10 pouce/harmonique). En pratique les formants sont rendus par un seul trait de pinceau qui couvre entièrement un canal et à moitié les deux canaux voisins. Toute ligne unie produit un son périodique. Pour reproduire les sons non périodiques, notamment les sifflantes et chuintantes on pointille aussi irrégulièrement que possible les zones du spectre sur lesquelles se concentre l'énergie. Les bruits d'explosions sont rendus par des traits brefs, plus ou moins verticaux, d'une largeur de bande d'environ 600 cps (5 canaux contigus).

Mue à la main ou entraînée par un moteur auxiliaire, la bande ainsi préparée défile devant deux cellules photoélectriques et sert de programme à la synthèse. Lorsque les structures sont peintes, la lumière réfléchie par les parties noircies de la bande de plastique tombe sur la cellule A. En cas d'utilisation à la place d'une bande de « visible speech » dessinée en négatif, d'une bande formée par un positif transparent, la lumière qui la traverse attaque la cellule B disposée de l'autre côté (Bibl. 10). Après amplification le courant issu des cellules est dirigé sur un haut-parleur et la voix est ainsi reproduite.

Ce système qui fait usage d'un semi-code a donné de bons résultats. L'appareil a l'avantage d'être très maniable et de pouvoir reproduire plusieurs spectrogrammes placés à la suite l'un de l'autre (10 sec. de discours env.). La préparation du programme de synthèse ne demande que très peu de temps et le dispositif est particulièrement intéressant en ce sens que les structures acoustiques mises en évidence par l'analyse peuvent être reproduites avec le nombre de détails que l'on désire et qui plus est, modifiées à volonté. Comme le programme de la synthèse ne comporte cependant, en-dehors du dessin en « code » des structures vocaliques et consonantiques qui apparaissent sur les spectrogrammes, aucune indication des variations mélodiques, le « pattern playback » qui est capable de produire une parole très intelligible si le sonagramme a été bien schématisé, présente hélas l'énorme inconvénient de ne pas pouvoir restituer les variations de la fréquence du ton laryngien : la voix produite est « recto-tono » sur un fondamental de 120 Hz, monotone et artificielle. La reproduction des sons de friction laissait également à désirer.

Le « PATTERN PLAYBACK » qui dans sa version primitive était donc sans grand intérêt pour les études portant sur l'intonation fut par contre un excellent instrument de re-



cherche pour l'étude de la *structure acoustique* des sons de leurs *transitions* et de leur *perception*.

F. S. COOPER, Philip LIEBERMAN et le regretté Pierre DELATTRE ont développé en effet au cours des quinze dernières années à partir des résultats qu'ils ont pu obtenir avec cet appareil, des idées extrêmement fécondes et déterminantes pour l'évolution des sciences phonétiques (19).

Des articles cités en référence, il ressort que sur la base de ces expérimentations P. DELATTRE fonda toute une « théorie » de la parole. Selon cette théorie, qui prise dans son ensemble peut être considérée comme une réponse à la question : « quels sont les éléments significatifs des enregistrements spectrographiques ? », les voyelles sont caractérisées par la fréquence et l'intensité de leurs formants (surtout des deux premiers), les consonnes par leur spectre propre et les transitions de formants des voyelles avoisinantes, ces transitions convergeant toutes vers une fréquence unique caractéristique de chaque consonne et dénommée « locus dental ».

#### *Le VOCODER PLAYBACK ou VOBACK.*

Au principe de « PATTERN PLAYBACK » se rattache aussi le « VOCODER PLAYBACK » ou VOBACK mis également au point par les Laboratoires HASKINS et décrit par L. O. SCHOTT dès 1948 dans un article intitulé : A Playback for visible Speech (Bell telephone Lab. Rec. 26, 1948 : p. 333-339). Tout comme le PLAYBACK cet appareil était basé en effet sur la lecture de sonagrammes peints, mais les structures stylisées servaient à contrôler un VOCODER à 18 canaux pour la reproduction de la voix. Deux lignes supplémentaires furent ajoutées en haut des spectrogrammes : la première compléta le programme de la synthèse en introduisant une distinction entre phonèmes voisés (sonores) et non voisés (sourdes) la seconde en lui fournissant un contrôle des variations de hauteur de la voix. Le VOBACK gardait ainsi les avantages essentiels du PLAYBACK (avec un peu moins de souplesse peut-être (20) du fait que le spectre fut réduit à 18 zones discrètes et que la source sonore devait passer de voix à friction dans tous les canaux en un même instant) mais il avait sur lui l'avantage de faire varier la hauteur de la voix, de reproduire d'une façon bien meilleure les fricatives et d'avoir beaucoup moins de bruit de fond. Il se prêtait ainsi non seulement à l'étude de la structure acoustique des sons mais aussi aux études portant sur l'intonation et l'accentuation. DELATTRE eut l'occasion de se servir des deux types d'appareils.

#### *L'ICOPHONE.*

Sur le principe qui consiste à reconvertir un enregistrement spectrographique en voix a été conçu aussi à une date plus récente (1956) un appareil élaboré par le Groupe de Re-

19) cf. à ce sujet quelques publications restées célèbres :

P. DELATTRE : Les indices acoustiques de la parole (Phonética vol II 1958) - de la hiérarchie des indices acoustiques pour la perception de la parole in *Proceedings of the Fifth International Congress of Phonetic Sciences* Münster 1964 p. 240 - 251 - Le jeu des transitions de formants et la perception des consonnes (in *Proceed. of the Fourth Inter. Cong. of Phonet. Sci.* Helsinki, 61 p. 407/17) - Some experiments on the perception of Synthetic Speech Sounds.

20) cf. F. S. COOPER - Speech synthesizers article cité p. 21.

cherche sur la Parole du Laboratoire d'Acoustique de la Faculté des Sciences de Paris (21) : l'ICOPHONE dont le schéma de principe est donné par la figure ci-dessous :

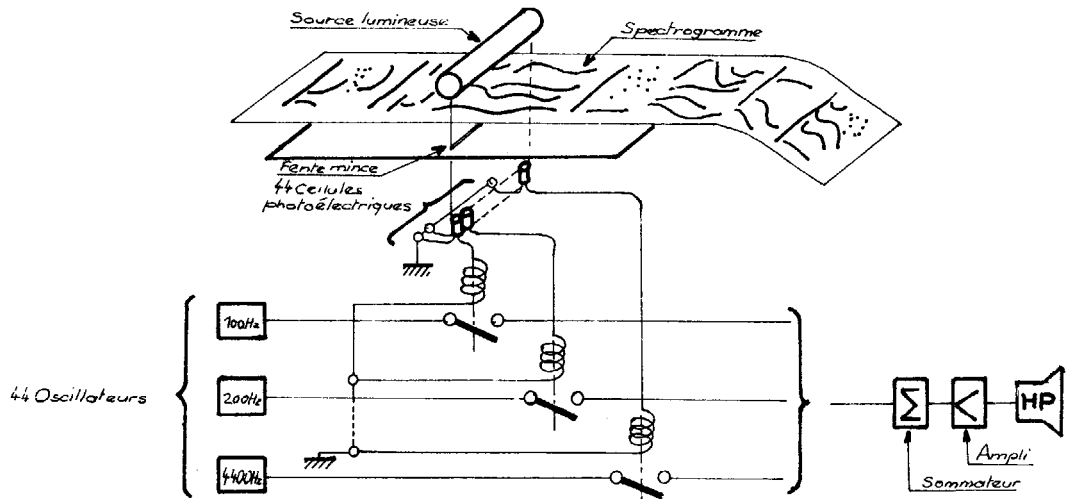


Schéma de principe de l'ICOPHONE (1956)  
du Groupe de Recherche sur la Parole du Laboratoire d'Acoustique  
de la Faculté des Sciences de Paris (Bibl. 18).

L'appareil décrit par J. S. LIENARD (Bibl. 18) présente tout comme le VOBACK un « VOCODER » pour la synthèse. Il s'agit en l'occurrence d'un ensemble de 44 générateurs autonomes disposés tous les 100 Hz et sur une gamme allant de 100 à 4400 Hz (ce qui est sans doute un peu juste si l'on songe que pour les sifflantes l'énergie se concentre sur des zones assez élevées et que le « pattern playback » allait jusqu'à 6 000 Hz). Ces oscillateurs sont pilotés par un nombre équivalent de cellules photoélectriques sensibles aux variations de la lumière engendrées par le déroulement devant une source lumineuse du sonagramme schématisé qui sert de programme à la synthèse. Lorsqu'une zone opaque des structures « stylisées » passe ainsi devant la cellule qui correspond à sa fréquence, cette cellule met en circuit un oscillateur. Les sons simples ainsi engendrés sont dirigés ensuite vers des circuits de mélange et d'amplification à la sortie desquels ils attaquent un haut-parleur. L'instrument permit à ses constructeurs d'obtenir « une voix chuchotée parfaitement naturelle » (Bibl. 18) et de « vérifier aussi un certain nombre d'hypothèses relatives à la structure de la parole ».

Le groupe reprit en fait, en les mathématisant ou mieux en les systématisant quelque peu et en les habillant de la Gestalt-théorie, les résultats de DELATTRE et l'idée d'un découpage du message parlé en sons ou plutôt unités acoustiques du langage. Ces « atomes » de langage constitués par l'assemblage de deux phonèmes reçurent le nom de « phonatomes ». Le nombre des « phonatomes » nécessaire dans la prononciation courante du français fut évalué à environ 900 unités. Comme certaines combinaisons sont extrêmement peu fréquentes, que d'autres sont réversibles et qu'à chaque « phonatome » correspond un « pattern »

21) Prof. SIESTRUNCK - E. LEIPP, Maître de Recherches au CNRS et M. CASTELLENGO.

bien précis on a établi qu'il suffirait en définitive d'un « stock » de 300 à 400 traces élémentaires pour synthétiser n'importe quel discours sans analyse préalable !

Cette tentative de simplification des données de l'analyse qui a abouti d'un côté à ce « microcosme » des « phonatomes » devait mettre en évidence d'un autre côté l'existence de structures essentielles, de paramètres généraux des sons du langage. La possibilité de reconstituer la parole à partir de constituants fondamentaux définis par une analyse préalable fut à l'origine d'une nouvelle génération d'appareils : *les synthétiseurs paramétriques*.

### SYNTHETISEURS PARAMETRIQUES.

Le « Parametric Artificial Talker » PAT) de W. LAWRENCE (22) et

L'« Orator Verbis Electrics » (OVE) de G. FANT (23).

La *synthèse paramétrique* qui recrée la parole au moyen des seuls *éléments significatifs* des sons (F0, F1, F2, F3, etc...) définis au préalable par une analyse spectrographique peut être considérée sans doute, par la qualité de la voix et des résultats qu'elle permet d'obtenir comme une des plus belles réussites des recherches qui aient été entreprises au cours de la seconde moitié de ce siècle dans le domaine de la parole artificielle, comme un succès aussi de l'électronique appliquée à la reproduction des sons du langage.

Plusieurs appareils qui offrent entre eux d'ailleurs de nombreuses ressemblances furent conçus sur le principe de la reproduction paramétrique : Le « *Parameter Artificial Talker* » construit par W. LAWRENCE en 1953 et connu plus familièrement sous le nom de PAT, l'« Orator Verbis Electrics » ou OVE mis au point par Gunnar FANT de l'Institut Royal de Technologie de Stockholm et le synthétiseur IBM en service à l'IBM Research Division de San Jose (California).

Dans sa version originale, PAT reformait la parole à partir d'un programme comportant six paramètres peints en lignes miniatures sur une plaque de verre explorée par un tube à rayons cathodiques. Il fut suivi de modèles plus perfectionnés, dotés d'un nombre plus élevé de lignes (8) et beaucoup plus maniables aussi qui fonctionnent à l'heure actuelle à Christchurch et à Edimbourg.

L'OVE I qui comportait un contrôle manuel de la fréquence du fondamental, de la source de sons vocaliques et de la fréquence des deux premiers formants fut suivi à son tour d'un modèle plus élaboré : l'OVE II mis au point par RISBERG, FANT et MARTONY, présenté à Oslo en 1958 au VIII International Congress of Linguistics et en 1961 à Philadelphie au Meeting de l'Acoustical Society of America (24). D'un intérêt évident pour les recherches de phonétique et d'acoustique, d'une maniabilité très grande, cet appareil présentait alors des performances si intéressantes et des possibilités si prometteuses que son principe fut retenu par la MELPAR, Inc. (25) pour la construction de l'« *Electrical Vocal Analog* » ou ÉVA, un des rares synthétiseurs qui ait été commercialisé, la plupart des autres étant toujours restés au stade de prototypes.

22) cf. W. LAWRENCE : The Synthesis of Speech from Signals which have a Low Information Rate. Communication Theory, ed. W. JACKSON (London 1953).

23) cf. G. FANT : Modern Instruments and Methods for Acoustic Studies of Speech. - Proceedings of the VIII International Congress of Linguistics (Oslo, 1958).

24) cf. à ce sujet le témoignage de F. S. COOPER in « Speech Synthesizers Proceedings of the Fourth International Congress of phonetic sciences - He'siak'i 1961 p. 3 - 13.

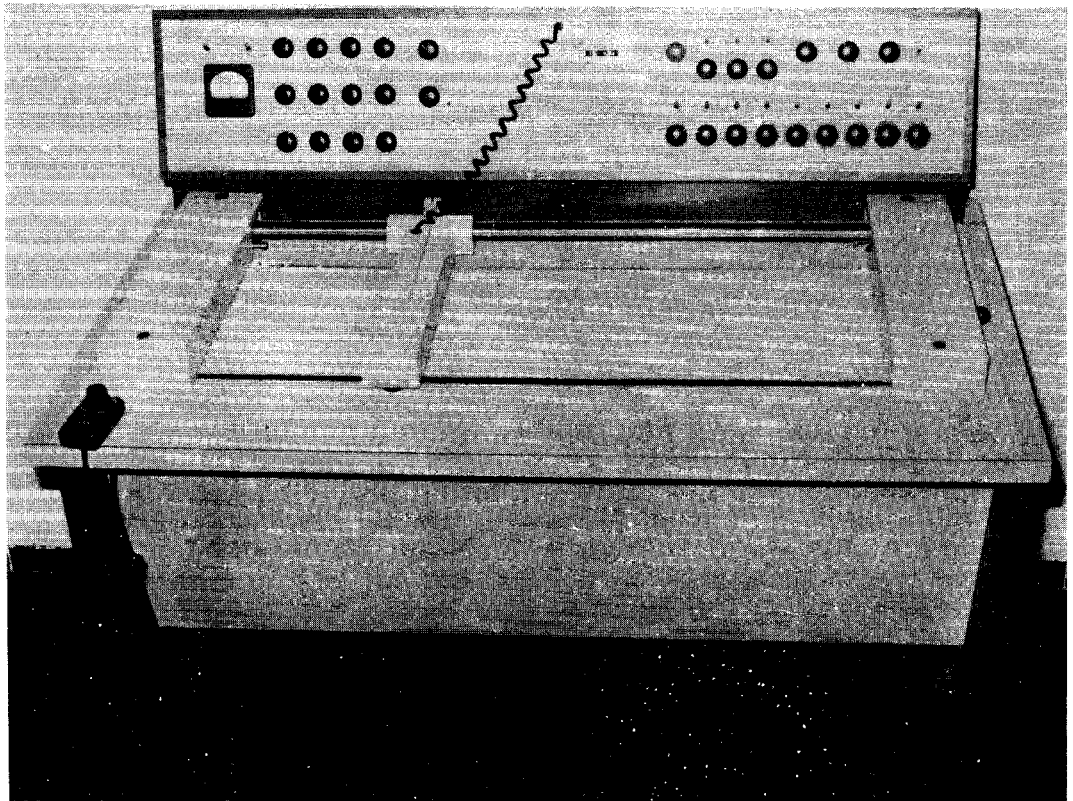
25) MELPAR, inc. Falls Church, Virginia 22046 - 7700 Arlington bld.

S'il est bien acquis aujourd'hui que les recherches phonétiques et toutes les sciences connexes qui s'appuient sur la connaissance que nous avons des sons du langage ne sauraient progresser sans le recours aux méthodes de synthèse qui permettent précisément d'étudier d'une manière approfondie les caractères significatifs de la parole, il faut bien constater d'un autre côté, que très peu de laboratoires possèdent jusqu'à ce jour de pareils appareils.

En acquérant dès 1967 l'EVA III, l'Institut de Linguistique et de Phonétique de l'Université d'Alger ne sacrifiait point à une mode mais à une nécessité impérieuse de la recherche et aux exigences d'un enseignement de haut niveau qui s'adresse chaque année à un nombre accru d'étudiants. En exprimant sa volonté délibérée d'entrer dans la compétition dans le domaine si important et si passionnant aussi des sciences liées à la parole, il accédait également au rang des Institutions les plus modernes qui sont appelées à devenir dans un avenir plus ou moins proche de vastes Centres du Langage...

*LE SYNTHETISEUR EVA III : COMPTE RENDU D'UNE SEANCE DE TRAVAIL ET  
ESQUISSE D'UNE METHODE*

*PRESENTATION DE L'APPAREIL : « Anatomie », caractéristiques techniques et fonctionnement.*



Le synthétiseur EVA III (fig. ci-dessus) se présente sous la forme d'un « meuble électronique » aux lignes très sobres. Il se compose d'une table munie d'un dispositif de lecture et destinée à recevoir la feuille de plastique, avec la représentation paramétrique qui sert de programme à la synthèse (Graphic Playback Unit), d'un châssis qui abrite les circuits élec-

ironiques qui à partir de cette représentation reproduisent la voix (speech synthesizer) et enfin d'un « tableau de bord » sur lequel se trouvent réunies toutes les commandes (Operational manifold chassis). Reprenons ces différentes parties plus en détail :

a) *Graphic playback Unit*

Le point de départ c'est toujours le tracé, l'« evagramme » qui sert de programme à la synthèse. Il s'agit d'une feuille de papier ou de celluloïd quadrillée en inches et dixièmes d'inch, d'environ 1 m de long sur 0,30 m de large (39 X 12 en inches) sur laquelle on représente au moyen d'une encre à base d'argent (26), spécialement conçue à cet effet et qui devient conductrice en séchant, les dix paramètres suivants de la séquence que l'on veut reproduire :

- 1) La fréquence du fondamental et ses variations (4<sup>e</sup> ligne en partant du haut du tracé, dénomination : F0 - Pitch),
- 2) La position relative, ou mieux les lignes d'évolution des quatre premiers formants. (lignes 2 à 8 en partant du haut du tracé - dénomination : F1, F2, F3, F4),
- 3) Les fréquences, zones de fréquences ou encore bandes de bruits qui définissent les spectres des fricatives, sifflantes (s, z) chuintantes (t, d) et phonèmes à turbulence effacée (f, v) (ligne 12, dénomination K1, Fricative pole),
- 4) L'amplitude de la voix pour tous les sons qui s'accompagnent de vibrations des cordes vocales (voyelles et consonnes sonores),
- 5) L'amplitude de bruit des phonèmes de la chaîne parlée. (Ligne 3, dénomination : Ac - Fricative Amplitude).

En fait, cette ligne sert à régler l'amplitude des bandes de bruits que l'on fait apparaître sur K1 ; on peut donc y faire figurer à la fois les variations d'intensité des constricatives et le bruit de l'explosion des occlusives.

- 6) L'amplitude de bruit pour le système de filtres vocaliques (production de la voix chuchotée). (Ligne 2, dénomination : AH - Aspiration Amplitude),

Cette ligne sert également à la production du bruit caractéristique de l'explosion des occlusives.

- 7) L'amplitude des composantes nasales dont les fréquences peuvent être réglées manuellement. (Ligne 1, dénomination : AN - Nasal Amplitude).

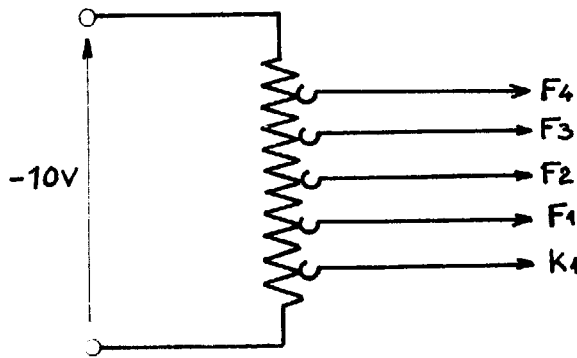
26) Silver préparation Electronic Grade - 4817 (Butyl Acetate Solution) E. I. DuPont de Nemours & Co. Electrochemicals Department - Wilmington, Delaware.

A chacune de ces lignes correspondent sur la feuille une position bien déterminée, des maxima et des minima entre lesquels sont comprises leurs fluctuations. Leurs coordonnées sont données par le tableau ci-après :

Etalonnage du programme de SYNTHESE						
Numéro de la ligne à partir du haut	Paramètre	MINIMUM		MAXIMUM		ECHELLE Nombre de cycles ou de volts par 1/10e d'inch.
		Distance de la ligne au bas de la feuille. (inch.)	Fréquence ou intensité correspondante	Distance au bas de la feuille	Fréquence ou intensité correspondante	
12	K1	0,8	3000 cps	1,2	7000 cps	1000 cps
11	F1	1,0	0 cps	2,0	1000 cps	100 cps
10	F2	1,5	500 cps	3,5	2500 cps	100 cps
9	F3	2,5	1500 cps	4,5	3500 cps	100 cps
8	F4	3,5	2500 cps	5,5	4000 cps	100 cps
7	ligne	non	utilisée	—	—	—
6	ligne	non	utilisée	—	—	—
5	AO	6,8	- 50 db	7,2	0 db	1 volt
4	FO	7,0	0 cps	9,0	400 cps	20 cps
3	Ac	8,5	- 50 db	9,0	0 db	1 volt
2	AH	9,3	- 50 db	9,7	0 db	1 volt
1	AN	11	- 50 db	11,4	0 db	1 volt

Ces données qu'il convient de respecter scrupuleusement si on veut obtenir de bons résultats permettent d'étalonner aisément le programme : à chaque ligne correspond une échelle précise, graduée en Hz lorsqu'il s'agit de variations de fréquence (lignes K1, F0, F1, F2, F3, F4), en décibels (db) pour les variations d'intensité (lignes Ao, Ac, Ah, An). (fig. Pl. 5).

Le dispositif de relecture (Chart Reader) est formé d'un bras mobile, ou mieux d'un chariot (movable carriage) coulissant le long d'un tube d'acier et entraîné par un moteur logé sous la table fig. p. 36) au moyen d'un système de transmission par câble et poulies. A ce chariot qui est muni en avant d'une roue sont suspendues deux résistances bobinées qui « balayent » le programme de la synthèse lorsque le mobile se met en mouvement. La première de ces bobines est en contact avec les cinq lignes du bas de la feuille (K1, F1, F2, F3, F4), la seconde avec les cinq autres (Ao, Fo, Ac, Ah, An) et chacune de ces résistances délivre sur les paramètres qu'elle couvre une tension qui varie d'un bout à l'autre du rouleau de 0 à 10 volts. Les deux ensembles ainsi constitués peuvent être assimilés chacun à un potentiomètre qui comporterait cinq curseurs et que l'on pourrait représenter schématiquement par la figure suivante :



Comme les lignes du programme sont conductrices d'électricité, elles recueillent en effet à chaque instant, à la manière des curseurs d'un potentiomètre, la tension qui correspond à leur point de contact avec les bobines. Pour un curseur ou un graphe donné, cette tension est évidemment fonction de la position du curseur le long de la résistance bobinée ou du degré d'élévation du graphe conducteur par rapport à sa ligne de référence. Les positions relatives des différentes lignes ainsi que leurs excursions maximales sont fixées par le constructeur : aux minima et maxima de tension correspondent les valeurs extrêmes évaluées en Hertz (paramètres fréquentiels) ou en déci-bels (paramètres d'amplitude) du tableau ci-dessus. Les tensions recueillies par les différentes lignes sont acheminées ensuite vers des circuits comparateurs où elles sont confrontées à des tensions de référence données par les fem. d'une série de piles (type Mallory Tr. 177 - 9. 8 volts).

A l'issue de cette opération, ou mieux, de ces circuits, les maximum et minimum de tension correspondent pour chaque ligne respectivement à 0 volt (position de repos) et - 4 volts (excursion maximale), conformément au tableau de calibration ci-après :

PARAMETRES	Tension (- volts)	FREQUENCES (Hz) OU AMPL. (db) CORRESP		LARGEUR DE BANDE	
		val. idéales	val. réelles	val. idéales	val. réelles
Restitution des spectres de bruits K1	0	3000 Hz	3000 Hz		480 Hz
	1	4000	3850		800
	2	5000	4950		1000
	3	6000	6200		1500
	4	7000	7500		1600
Ligne d'évolution du 1er formant F1	0	0 Hz	120 Hz	70 Hz	70 Hz
	1	250	250	70	70
	2	500	515	70	65
	3	750	750	70	70
	4	1000	1000	70	70
Ligne d'évolution du 2e formant F2	0	500 Hz	500 Hz	80 Hz	80 Hz
	1	1000	930	80	65
	2	1500	1400	80	80
	3	2000	1935	80	80
	4	2500	2500	80	80
Ligne d'évolution du 3e formant F3	0	1500 Hz	1500 Hz	100 Hz	100 Hz
	1	2000	1880	100	100
	2	2500	2400	100	115
	3	3000	2975	100	125
	4	3500	3500	100	140

PARAMETRES	Tension <sup>e</sup> (- volts)	FREQUENCES (Hz) OU AMPL. (db) CORRESP.		LARGEUR DE BANDE	
		val. idéales	val. réelles	val. idéales	val. réelles
Ligne d'évolution du 4 <sup>e</sup> formant F4	0	2500 Hz	2500 Hz	140 Hz	150 Hz
	1	3000	2900	140	150
	2	3500	3360	140	150
	3	4000	4000	140	150
	4	4500	4500	140	150
Amplitude vocale (voyelles et conson- nes sonores) Ao	4,0	0 - db	0 - db		
	1,26	10	10		
	0,4	20	20		
	0,126	30	34		
	0,04	40	48		
Ligne d'évolution du fondamental - (ligne mélodique) Fo	0	0 Hz	20 Hz		
	1	100	100		
	2	200	195		
	3	300	295		
	4	400	395		
Amplitude des bandes de bruits de la ligne K1 Ac	4,0	0 - db	0 - db		
	1,26	10	10		
	0,4	20	20		
	0,126	30	34		
	0,04	40	48		
Amplitude de la voix pour le système de filtres vocaliques AH	4,0	0 - db	0 - db		
	1,26	10	10		
	0,4	20	20		
	0,126	30	34		
	0,04	40	48		
Amplitude des composantes nasales AN	4,0	0 - db	0 - db		
	1,26	10	10		
	0,4	20	20		
	0,126	30	34		
	0,04	40	48		

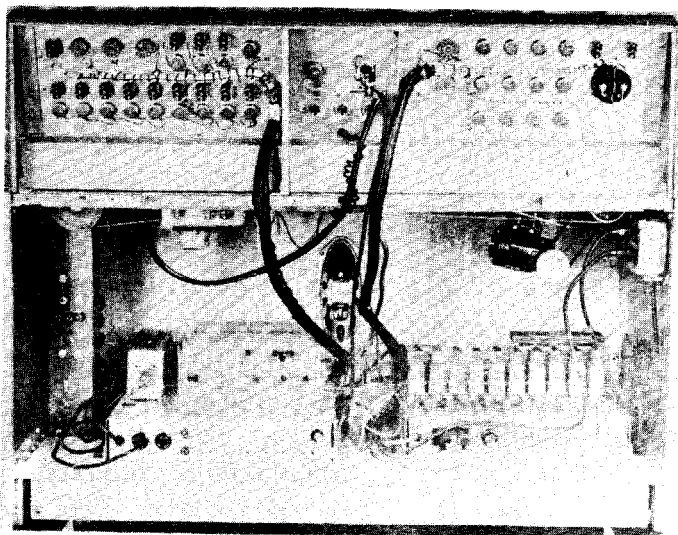
Le dispositif de lecture convertit ainsi les graphes des *dix paramètres* en un ensemble de signaux qui actionnent les circuits électroniques équivalents du synthétiseur de la parole.

- Les *fréquences* et les *amplitudes* des composantes de la chaîne parlée portées en *ordonnée* sur le tracé sont rendues par les « excursions » latérales, les *déplacements verticaux* des lignes conductrices le long des résistances bobinées.
- le déroulement du discours dans le *temps* porté en *abscisse*, par le *déplacement horizontal* du chariot dont la *vitesse* (qui correspond par conséquent au « débit ») est normalement de 2  $\frac{1}{2}$  inches par seconde et réglable à volonté par la rotation d'un bouton de commande.

b) *Le châssis et les circuits électroniques équivalents (Speech synthesizer).*

Le châssis renferme un ensemble de circuits électroniques qui, commandés par le programme de synthèse, ont pour rôle de reconstituer la voix (fig. ci-après).





Le Synthétiseur EVA III (Anatomie de l'appareil)

Dans la mesure où cet ensemble est capable d'engendrer tous les sons du langage et de restituer la parole, l'EVA III peut être considéré comme un « équivalent » de l'appareil phonatoire humain, conformément à sa dénomination, mais il ne doit pas être confondu en cela avec d'autres synthétiseurs du type de la LEA (Line Electrical Analog) également mis au point par FANT et sur lesquels on a essayé de reproduire au moyen de circuits électroniques, et en le découpant en un nombre sans cesse croissant de « tranches » successives, le tractus humain. Le schéma d'ensemble de l'EVA III dérive évidemment de celui de l'OVE II donné par la fig. ci-dessous :

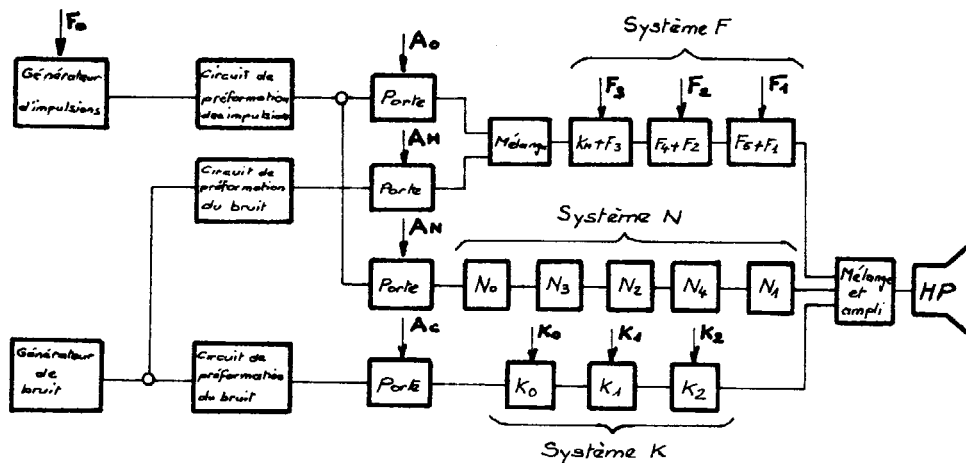


Schéma du synthétiseur OVE II de Stockholm

(FANT et MARTONY 1962)

L'appareil fonctionne sur un secteur alternatif (50-60 Hz, 115 volts) et sa consommation maximum est de 5 Amp. Il se compose pour l'essentiel :

- d'une alimentation stabilisée (fig. page 31),
- d'une série de trois circuits imprimés (fig. page 31) qui en constituent les pièces maîtresses et qui sont les « équivalents » du modèle humain,
- et d'un amplificateur de sortie avec un haut-parleur qui transforme les oscillations électriques en vibrations sonores (fig. page 31).

— Le premier des trois circuits sert à l'excitation. Il se compose d'un générateur d'impulsions et d'un générateur de bruits et joue pour les deux autres « équivalents » tantôt le rôle des cordes vocales et de la fourniture glottique *initiale* qui fait entrer en résonance le pavillon pharyngo-buccal lors de la production des phonèmes sonores, tantôt le rôle du souffle qui engendre sur son passage des bruits de friction (production des consonnes sourdes) et parfois même les deux à la fois : c'est le cas par exemple lorsqu'il s'agit de produire une fricative sonore pour laquelle la composante laryngée vient se superposer à un spectre de bruits. Le générateur d'impulsions est commandé par la ligne F0 ou manuellement par la rotation d'un bouton : la fréquence est susceptible de varier de 20 à 400 Hz.

— La deuxième carte se compose des *circuits électroniques* nécessaires à la production des *nasales* et des *spectres de bruits*. Le « murmure nasal » du segment occlusif des consonnes est fourni par un équivalent électrique du modèle humain sous la forme d'un spectre comprenant *trois résonances* et une *anti-résonance*, dont on peut faire varier manuellement les valeurs au moyen d'une série de commandes, conformément au tableau ci-dessous :

« NASAL ANALOG »						
Anti-résonance (zéro)	N0	0 - volts	600 Hz	600 Hz	200 Hz	120 Hz
		1	1050	1000	200	120
		2	1500	1400	200	135
		3	1950	1875	200	120
		4	2400	2400	200	220
Résonance 1 (pole)	N1	1	100 Hz	100 Hz	200 Hz	150 Hz
		2	175	165	200	186
		3	250	250	200	191
		4	325	320	200	180
		5	400	385	200	200
Résonance 2 (pole)	N2	0 - volts	400 Hz	400 Hz	200 Hz	185 Hz
		1	700	670	200	200
		2	1000	1000	200	200
		3	1300	1330	200	210
		4	1600	1620	200	250
Résonance 3 (pole)	N3	1	600 Hz	600 Hz	180 Hz	175 Hz
		2	1500	1050	180	175
		3	1500	1460	180	175
		4	1950	1930	180	180
		5	2400	2400	180	220

Les mêmes circuits peuvent servir à donner aux voyelles nasales leur coloration particulières (27). L'amplitude des composantes nasales est déterminée dans tous les cas par les variations de ligne AN.

27) Cette question est assez complexe. Une étude d'ensemble est en préparation à ce sujet. Elle doit paraître prochainement.

Les circuits de formation des bruits qui figurent sur la même carte sont commandés par la ligne 12 et permettent d'engendrer toutes les fricatives au moyen d'une bande de fréquences assez large pouvant se situer entre 3 000 et 7 000 Hz (Fricative pole K1) et d'un « zéro » (Fricative zéro K0) dont la position peut varier entre 750 et 1 750 Hz, conformément au tableau ci-après :

« FRICATIVE ANALOG »						
COMPOSANTES SPECTRALES		Tension (volts)	FREQUENCES (Hz cor.)		LARGEUR DE BANDE	
			val. idéales	val. réelles	val. idéales	val. réelles
anti-formant (fricative zéro)	K1	0	750 Hz	840 Hz		90 Hz
		1	1000	1100	---	110
		2	1250	1360	non	175
		3	1500	1750		240
		4	1750	2100		350
bande de fréquences (fricative pole)	K0	0	3000 Hz	3000 Hz		480 Hz
		1	4000	3850		800
		2	5000	4950	spécifiées	1000
		3	6000	6200		1500
		4	7000	7500		1600

Rappelons à ce sujet que les maxima de densité spectrale se situent pour les sifflantes au-dessus de 4 000 Hz, dans les régions supérieures du spectre et sensiblement plus bas pour les chuintantes (deux zones généralement comprises entre 1 500 et 4 500 Hz) (28). L'amplitude de ces bandes de bruits est déterminée, comme dit, par Ac.

La troisième carte se compose des circuits nécessaires à la production des sons vocaliques. Ce « Vocal Analog » permet d'obtenir des spectres constitués de cinq zones formantiques dont les fréquences respectives sont déterminées par les lignes F1, F2, F3, F4 (se reporter au tableau de la page 34) et une commande manuelle F5 dont les valeurs extrêmes sont 3 000 et 5 000 Hz. Tous les timbres vocaliques peuvent être engendrés. La largeur des différents formants est en moyenne de 70 cps pour le premier, de 80 pour le suivant, de 100 cps pour le troisième, de 150 et 400 cps pour les deux derniers. Les amplitudes relatives des diverses composantes sont fixées par le constructeur et non susceptibles (hélas !) d'être modifiées. La carte comporte en plus un circuit KH (Higher Pole correction) dont le rôle est de restituer dans une certaine mesure les fréquences supérieures des spectres. Les caractéristiques sont les suivantes :

KH	Fréquence (Hz)		Largeur de bande	
	idéale	réelle	idéale	réelle
	4 500 Hz	4 700 Hz	4 500 Hz	4 500 Hz

28. cf. PETER STREVEN'S Spectra of Fricative Noise in Human Speech in Language and Speech vol. 3 (1960) p. 32 - 49 ainsi que notre mémoire de maîtrise : La méthode sonographique en phonétique expérimentale (Fac - Lettres - Strasbourg).

Le schéma d'ensemble de l'appareil est donné par la figure de la planche n° 1. Son fonctionnement est aisé à comprendre :

Prenons par exemple le cas d'une voyelle orale. Il faudra, pour produire ce phonème, disposer d'impulsions à une certaine fréquence ; ce rôle incombera à l'oscillateur de relaxation que commande F0. Il suffira ensuite d'ouvrir la porte A0 qui donne accès aux circuits des voyelles orales lesquels imprimeront alors au signal une forme qui sera fonction des positions respectives des lignes F1, F2, F3, F4 du tracé et de la position de la commande manuelle F5. Lorsqu'il s'agit de produire une voyelle nasale, on ouvre en plus, au moyen de la ligne AN la porte qui donne accès aux circuits équivalents du conduit nasal de la deuxième carte. Le « murmure nasal » s'ajoute alors à la structure en formants qui donne à la voyelle son timbre propre et lui confère la coloration particulière des phonèmes qui se prononcent voile abaissé.

Supposons maintenant que nous ayons à reproduire une fricative sourde, la sifflante [S] par exemple. L'énergie sera fournie dans ce cas par le générateur de bruits et l'amplitude du signal sera déterminée par la ligne AC qui commande la porte donnant accès aux circuits équivalents de la carte 2. La délimitation en fréquence de la zone de bruits requise se fera au moyen de la ligne K1 qui commande ces circuits et qui peut varier entre 3 000 et 7 000 Hz. Dans le cas où nous nous sommes placés, cette ligne devra être à son maximum, le spectre du [S] comportant des fréquences très élevées.

Le processus est le même lorsqu'il s'agit d'engendrer une *fricative sonore* mais on ajoute dans ce cas au bruit la composante laryngée qui s'y superpose normalement. Elle sera fournie par le générateur d'impulsions à la fréquence F0 désirée et son amplitude sera déterminée au moyen de la ligne A0 qui commande la porte donnant accès au « Vocal Analog ». Pour les sourdes cette porte doit être complètement fermée (position de repos de la ligne). Les composantes harmoniques étant généralement de faible intensité, il suffira de l'ouvrir très légèrement pour la formation des sonores sifflantes et chuintantes, davantage par contre pour des phonèmes qui, par leur constitution acoustique, sont très proches des voyelles comme c'est le cas des liquides [l] et [r], très largement pour les semi-consonnes (j, y, w) dont les spectres sont très peu chargés des bruits. De l'ouverture de la porte commandée par A0 dépend en quelque sorte le *degré de vocalisation* des phonèmes, du signal issu de Ac, le degré de friction ou de spirantisation. Les amplitudes respectives de ces signaux peuvent être « dosées » à volonté.

Les *occlusives* bilabiales (p, b, m), alvéo-dentales (t, d, n), vélaires (k, g) requièrent pour leur phase audible (explosion) de brusques variations d'énergie. Les « barres de bruits » produits lors de la rupture de l'occlusion par l'échappement très brusque de l'air phonateur peuvent être rendus à la synthèse à la fois au moyen de la ligne AH qui commande la porte d'entrée du signal à spectre continu vers le « Vocal Analog » et qui sert également à la production de la voix chuchotée et de la ligne Ac qui commande la porte d'entrée du même signal vers les circuits formateurs des spectres des fricatives. Les transitions de formants (F2 en particulier) servent à la perception de leur lieu d'articulation (29). Les lignes A0 et AN déterminent comme précédemment la sonorité et la nasalité des phonèmes.

29) cf. DELATTRE. Le jeu des transitions de formants et la perception des consonnes in Proceedings of the Fourth International Congress of Phonetic Sciences Helsinki 1961 (The Hague, Mouton & Co 1962) p. 407 - 417.

c) *Le tableau de commande (operational manifold chassis).*

Le « tableau de bord » de l'appareil contient sur le côté gauche toute une série de boutons destinés au réglage du niveau des lignes de référence des différents paramètres, ainsi qu'un indicateur de contrôle qui permet de réaliser les ajustements nécessaires. Le côté droit comporte par contre tout un ensemble de leviers et de boutons de commande ; ils permettent de faire fonctionner l'appareil soit en position automatique, les circuits étant actionnés par les signaux issus des différents paramètres, soit en position manuelle les circuits de synthèse étant commandés dans ce cas par la rotation de toute une série de boutons prévus à cet effet. Le bras de lecture peut être également désenclenché du système de transmission et déplacé à la main.

Essayons de réaliser maintenant, à titre d'exemple, le programme de synthèse d'une séquence.

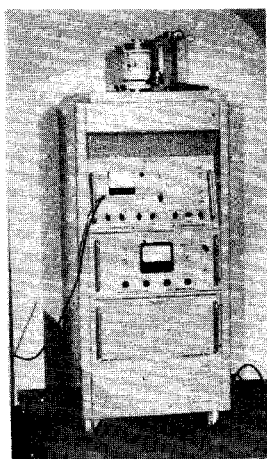
« Cette voix est le résultat d'une découverte scientifique ».  
Réalisation d'un « évagramme ».

Pour accentuer le côté spectaculaire et l'étonnement que l'on ressent toujours en présence d'essais de voix artificielle, pour célébrer aussi cette victoire de la science qui est arrivée à conférer le don de parole à une machine, pour les besoins enfin de la démonstration, nous avons retenu la séquence suivante :

« Cette voix est le résultat d'une découverte scientifique »  
en transcription phonétique (Notation internationale) :

I Sɛt vwa ɛ la Rezylta dynə dekuvɛRtə Sɔatifik(ə) I

La préparation du programme de synthèse se fait en trois temps : le premier consiste à enregistrer et à analyser la séquence, le second est nécessaire à la reproduction sur papier des différents paramètres, le dernier est réservé enfin aux « retouches » et à la mise au point définitive du tracé.

a) *Enregistrement et analyse sonographique de la séquence.*

La séquence que l'on veut produire artificiellement doit, de préférence, être enregistrée tout d'abord sur bande magnétique. La parole synthétique pourra être comparée ainsi en fin d'opération à la voix humaine et dans l'immédiat cet enregistrement servira à l'analyse sonographique de la séquence.

Le sonographe (30) ne peut analyser de toute manière que des « portions » de discours de 2 s 4/10 ; on est donc obligé de procéder par enregistrements successifs et de recoller les tracés obtenus. Ces tracés sont réalisés en filtrage large (300 cps) tout d'abord sur l'échelle normale (85-8 000 Hz ; pl. 6) puis sur l'échelle dilatée (85-4 000 Hz ; fig. pl. 3). Sur cette même échelle on réalise ensuite les sonagrammes en filtrage étroit (fig. pl. 4) en y faisant figurer aussi au moyen du dispositif « Amplitude display » les variations de l'intensité.

b) *Analyse des résultats sonographiques et reproduction sur papier des différents paramètres.*

Les tracés sonographiques une fois réalisés, les phonèmes de la séquence à reproduire doivent être très soigneusement délimités. Les mêmes divisions qui correspondent à la durée des sons et au déroulement du discours dans le temps sont reportées ensuite sur le programme de la synthèse au moyen de longs traits verticaux.

À ces dispositions préliminaires fait suite alors un travail beaucoup plus long qui consiste en un examen minutieux des tracés fournis par l'analyse sonographique et un « *recensement* » des caractères et indices acoustiques des différents phonèmes de la séquence considérée. Au cours de cette opération qui s'accompagne tout normalement d'une réflexion sur la *genèse des sons et leur mode de production*, les différents phonèmes sont analysés à travers la « grille » des paramètres qui commandent le synthétiseur de la parole. En remplissant les cases correspondantes aux différents indices acoustiques des phonèmes ou obtient un tableau du type de celui de la pl. 2, une « *matrice de conversion* » en quelque sorte qui permet de passer très aisément par la suite des données de l'analyse sonographique à la *représentation paramétrique*, « stylisée », des sons du langage.

À partir de là, la réalisation du tracé devient, en effet, un jeu d'enfant. Les lignes d'évolution des quatre premiers formants sont décalquées directement des sonagrammes en filtrage large (échelle dilatée - 85 - 4 000 Hz) sur un papier transparent ou une feuille de plastique coupée aux dimensions du tracé sonographique et reportées ensuite sur l'évagramme (pl. 3). La ligne qui correspond à la fréquence 1 000 Hz sert généralement de référence. Le tracé sonographique doit donc être nécessairement gradué ou étalonné du moins à l'aide d'un générateur de fréquences. L'opération qui vient d'être décrite peut soulever quelques difficultés dans la mesure où les transitions de formants n'apparaissent pas toujours avec une entière évidence, dans la mesure aussi où le bas des tracés est généralement trop noir et le haut trop pâle pour qu'on puisse discerner les structures avec une entière certitude. On peut remédier dans une certaine mesure à ces insuffisances en réduisant au moyen du « mark level » l'intensité de la trace lorsque le stylet est au début de sa course et en l'augmentant progressivement au fur et à mesure qu'il s'élève.

La ligne mélodique F0 s'obtient en décalquant sur le sonagramme réalisé en filtrage étroit (échelle dilatée, 85 - 4 000 Hz) la ligne d'évolution du 5<sup>e</sup> harmonique. Les variations de hauteur qui n'apparaissent pas suffisamment au bas des enregistrements sont ainsi multipliées par cinq et nettement visibles (pl. 4). La ligne qui correspond à la fréquence de 200 Hz sert généralement de référence pour le passage du sonagramme à l'évagramme. Ce tracé doit donc être également étalonné.

Les lignes A0 et Ac qui déterminent respectivement l'amplitude de la voix (voyelles et consonnes sonores) et l'amplitude des bruits (fricatives et occlusives, sourdes et sonores) s'obtiennent à partir du graphe de l'évolution de l'intensité globale donnée par « l'amplitude display » de l'analyseur (pl. 4).

Les lignes KI (fréquences des spectres de bruits) AH (Aspiration amplitude) et AN (Amplitude des composantes nasales) sont construites en fonction des données de l'analyse sonographique (pl.4). La ligne AH se superpose en quelque sorte à la ligne Ac pour la production de l'explosion des occlusives.

Tous les graphes sont réalisés d'abord au crayon, puis à l'encre spécialement destinée à cet usage. On vérifie, lorsqu'elle est sèche, au moyen d'un ohmmètre que les différentes lignes sont bien conductrices d'électricité. Le tracé est alors prêt pour la synthèse (Bibl. 5).

c) « Cette voix est le résultat d'une découverte scientifique » Audition de la séquence et « retouches ».

Le moment le plus attendu et le plus exaltant aussi est évidemment celui où le tracé ainsi réalisé sert pour la première fois. Si le travail est fait avec soin et précision les résultats sont excellents. Tous ceux, amis et collègues, qui ont pu entendre la séquence que nous venons d'analyser pourraient le confirmer. Les tracés sont d'ailleurs là pour le prouver ; la parole artificielle produite par le synthétiseur a été enregistrée à son tour sur bande magnétique pour pouvoir être comparée à tout moment à la voix humaine et la séquence a été analysée aussi au sonographe pour pouvoir être confrontée à l'original (pl. 6). Les tracés paraissent un peu différents du fait que l'intensité des formants ne peut pas être réglée sur la machine mais pour le reste ils sont profondément ressemblants et à l'audition la phrase est parfaitement intelligible.

Lorsqu'un passage ne donne point entière satisfaction, il est possible d'ailleurs d'y apporter quelques « retouches ». Il suffit de comparer pour cela l'enregistrement sonographique de la séquence artificielle à l'original, de déceler les éventuels écarts et de rectifier en conséquence le tracé des paramètres : l'encre s'enlève très facilement à l'aide d'un bouchon de coton imbibé d'acétate de butyl.

Ajoutons que le fait d'entendre parler une machine produit toujours une impression très vive sur les auditeurs, surtout lorsqu'il n'ont jamais eu l'occasion d'assister à une pareille tentative ; nous en avons souvent fait l'expérience en début d'année avec nos étudiants du certificat de phonétique qui sont d'ailleurs toujours très enthousiasmés lorsque nous leur disons qu'ils sauront, au bout de quelques mois, en faire autant. De l'étonnement à l'admiration, les sentiments les plus divers naissent et se confondent généralement en cet instant où se défait l'un des mythes les plus anciens de l'humanité...

La préparation d'un programme de synthèse, aussi intéressante soit-elle, ne constitue à vrai dire que très rarement un but en soi. Elle n'est le plus souvent qu'une étape préliminaire de travaux bien plus importants ; la véritable recherche ne commence généralement que lorsque le tracé est au point. Il est en effet possible alors de *modifier la valeur des paramètres* voire même de les supprimer comme on l'entend et de *faire subir à la phrase de synthèse toutes les transformations* que l'on désire (31). En faisant varier par exemple la fréquence du fondamental, une phrase énonciative peut devenir interrogative ou exclamative, traduire une affirmation catégorique ou, au contraire l'étonnement ou le doute (Bibl. 86). Il

31) à l'exception toutefois de l'intensité des formants dont la valeur reste constante.

est possible aussi de modifier l'un après l'autre tous les paramètres et de voir dans quelle mesure la phrase demeure intelligible. Par sa *très grande maniabilité*, l'appareil se prête aux recherches les plus diverses et dont certaines débouchent d'ailleurs sur des *applications pratiques* (prothèse pour sourds, matériel de phono-audiologie, répondeurs automatiques etc...). Il convient tout autant aux recherches spécialisées de phonétique acoustique qu'aux études relatives à *l'intelligibilité* de la parole, à la perception ou encore à l'étude des *faits d'accents et de mélodie*. La place nous manque ici pour traiter d'une manière suivie de tout l'éventail des possibilités qu'il offre aux chercheurs ; aussi nous contenterons-nous de regrouper sous la forme d'*un tableau* plus indicatif qu'exhaustif, un certain nombre d'*applications* dont la méthode est susceptible. Les *trois volets* dont il se compose correspondent aux différentes orientations de la recherche et comportent en outre un certain nombre de renvois à des travaux déjà effectués dont on trouvera les références dans notre bibliographie.

### APPLICATIONS : a) EN PHONÉTIQUE ACOUSTIQUE

#### *Etude de la structure acoustique des sons du langage*

- Signification acoustique des différentes composantes spectrales des sons du langage. (zones formantiques, bandes de bruits, importance de la composante laryngée pour les fricatives sonores, importance des bruits etc...),
- Approfondissement de la connaissance des résultats obtenus à l'analyse,
- Synthèse des sons vocaliques et consonantiques et étude de leur mode de production (Bibl. 51).
  - *Synthèse des voyelles* : Détermination de la *loi de formation des timbres* (marche des formants, positions relatives des différents formants étude de leurs variations etc...).
  - *Synthèse des sons consonantiques* : Mode de production des différentes constrictives (sifflantes, chuintantes, liquides l, r et leurs variantes) fréquence, largeur, intensité des bandes de friction.
    - Importance du voisement par rapport à la structure en bruits.
    - Mode de production des semi-consonnes (caractères distinctifs de ces phonèmes)
    - Mode de production des occlusives (Bibl. 54).

#### *Etude de l'enchaînement des structures sonores dans la chaîne parlée*

- Etude des *transitions* de formants et approfondissement de la notion de *locus dental* (Bibl. 48).
- Influences mutuelles des articulations dans la chaîne parlée (Bibl. 56).

#### *Voix humaine et voix artificielle* (Bibl. 59).

- Caractères vocaliques et extra-vocaliques des sons du langage (timbre vocal  $\neq$  timbre vocalique).
- Etude des composantes individuelles.
- Essais de « personnalisation » de la voix.



## SCIENCES ET DISCIPLINES CONNEXES

*Techniques de traitement du langage parlé*  
(mise au point des procédés de reconnaissance automatique) (Bibl. 11,64).

*Transmissions et technique des télécommunications*  
(économie des lignes téléphoniques) compression de bande (Bibl. 60, 61, 62, 63, 67, 69).

*Automation*  
*Applications industrielles*  
la « dactylo automatique » réalisation d'unités de réponse vocale etc. (Bibl. 11, 18, 64, 68).

## APPLICATIONS : b) PERCEPTION ET INTELLIGIBILITE

*Détermination des caractéristiques physiques servant à l'identification des phonèmes d'une langue*

(indices de fréquence, d'intensité, de hauteur de durée etc...) (Bibl. 70, 71, 72, 73).

———— Limite des phonèmes (Bibl. 73).

———— Importance pour la compréhension de la parole des diverses zones formantiques (F1 et F2 en particulier).

———— Importance pour la perception des sons consonantiques des bandes de bruits et des transitions de formants.

———— L'Intonation et l'identification des mots dans la chaîne parlée (Bibl. 78).

*Hiérarchie des indices acoustiques pour la perception de la parole* (Bibl. 70).

*Degrés et limites de l'intelligibilité.*

———— Essais d'appauvrissement du spectre et

———— Détermination des « structures minimales » (Bibl. 79).

(composantes véhiculant l'information nécessaire et suffisante).

## SCIENCES ET DISCIPLINES CONNEXES

Audio - phonologie  
(prothèses pour sourds)  
(Bibl. 80, 81, 82, 83).

Transmissions et Technique des  
Télécommunications  
(Bibl. 83, 84).

## APPLICATIONS : c) ETUDE des FAITS d'ACCENT et de MELODIE

*Rôle de la FREQUENCE, de l'INTENSITE et de la DUREE dans la perception de l'ACCENT et de la MELODIE* (Bibl. 85).

(Etudes spéciales - comparatives - générales)

*Etude des facteurs ecto-sémantiques de l'intonation* (Bibl. 86).

———— caractérisation de ces facteurs par synthèse.

« patterns » des phrases - interrogative (?)

- énonciative (.)

- exclamative (!)

- dubitative etc...

- étude des *inflexions de la voix, de ses contours, des éléments affectifs* du langage (Bibl. 88) - expression de la colère  
- expression de la surprise  
- expression de la joie etc...
- détermination des « seuils différentiels » de perception de ces nuances (Bibl. 87) (passage de la phrase énonciative à la phrase interrogative etc...).
- *Détermination des faits de substrats* (caractérisation des différents parlers).

### SCIENCES ET DISCIPLINES CONNEXES

#### *Psychologie*

(dans la mesure où l'intonation est en rapport direct avec l'expression des idées et des sentiments)

#### *Linguistique*

(dans la mesure où « chaque langue articule ses messages selon un code particulier » Martinet)

#### *Socio - linguistique*

(dans la mesure où l'intonation est déterminée par l'usage et obéit à des règles spéciales à chaque communauté)

#### *Grammaire*

(dans la mesure où les facteurs les plus simples sont répertoriés par elle)

Si l'intérêt d'une méthode instrumentale s'évalue d'après le nombre et l'importance des études qu'elle est susceptible de servir, les mérites de l'EVA III - et des techniques de synthèse d'une manière générale - ne sont plus à démontrer.

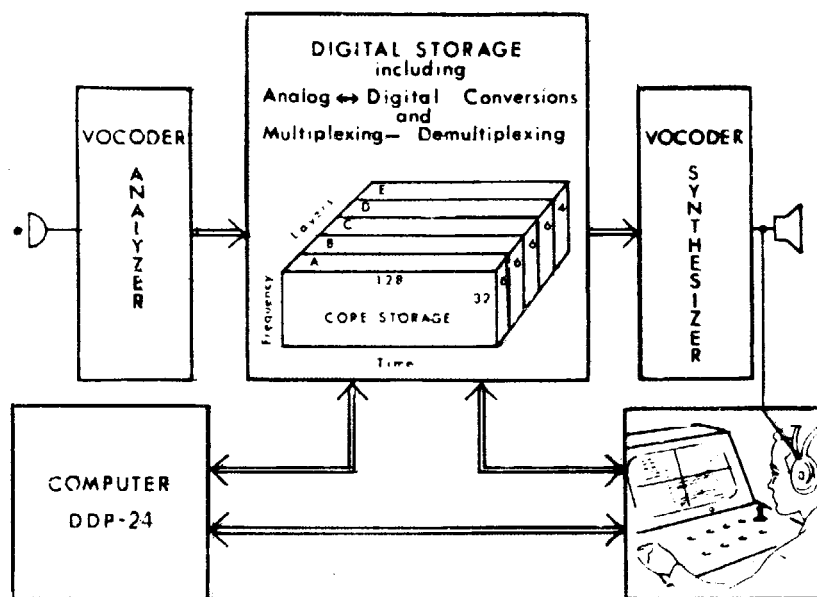
L'appareil constitue aujourd'hui l'âme même de notre laboratoire d'acoustique. Les étudiants du certificat de phonétique initiés aux mystères de la parole artificielle apprennent à s'en servir ; les chercheurs s'y relayent...

La méthode nous a déjà servi dans la détermination des attributs acoustiques de la nasalité, l'étude des lois de formation des timbres (variations de la position des formants, de F2 en fonction de F1, des rapports F2/F1, F3/F1, etc...) et bien d'autres recherches sont actuellement en cours.

Mais, si les sciences phonétiques et connexes ont pu, grâce à la parole artificielle, réaliser des progrès étonnants au cours des dernières années, la technique a évolué d'une manière plus surprenante encore, et alors que la plupart des laboratoires n'ont pas encore pu disposer jusqu'à ce jour des crédits suffisants qui leur auraient permis de se mettre à l'heure de la synthèse, elle produit chaque jour de nouveaux prototypes. Les derniers nés font tous appel à l'*ordinateur* ce « nouveau venu » qui a déjà révolutionné tant de disciplines. Divers emplois en ont déjà été faits jusqu'à ce jour.

— L'idée première qui vint à l'esprit des constructeurs fut de commander ou de programmer par ce moyen des Vocoders. Diverses tentatives ont été entreprises dans ce sens par des chercheurs comme J. L. FLANAGAN, C. H. COKER et CH. BIRD (1), J. L. GERSTMAN (2), C. LOCHBAUM (3) et V. A. VYSSOTSKY (4).

C'est d'une idée analogue que procède au ssi le Digital Spectrum Manipulator mis au point par les Laboratoires HASKINS et décrit par F. S. COOPER (5) au cinquième Congrès International des Sciences Phonétiques (fig. ci-dessus).



*Digital spectrum Manipulator (d'après F. S. COOPER)*

— Parallèlement aux synthétiseurs paramétriques se sont développés aussi les appareils analogiques du type de la LEA (Line Electrical Analog- de G. FANT (6), de la DAVO (Dynamic Vocal tract Analog) et de la DANA (Dynamic Nasal Analog) de ROSEN (7) STEVENS et DENNIS du M.I.T., qui permettent de simuler en quelque sorte le fonctionnement des organes phonateurs au moyen de circuits électriques et qui présentent de ce fait

- 1) J. L. FLANAGAN, C. H. COKER et CH. BIRD  
Computer Simulation of a Formant Vocoder synthesizer  
in J. Acoust. Soc. Am. 35, 2003 (A) 1962.
- 2) J. L. KELLY et L. J. GERSTMAN  
An Artificial Talker Driven from a Phonetic Input. J. Acoust. Soc. Am. 33, 835 (A) 1961.
- 3) C. LOCHBAUM  
Speech Synthesis Proc. Stockholm  
Speech Comm. Seminar Sweden, september 1962.
- 4) V. A. VYSSOTSKY - A Block Diagram Compiler. Bell System Tech. J. 40, 669 - 676 (1961).
- 5) F. S. COOPER Speech Synthesizers - in Proceedings of the Fourth International Congress of Phonetic Sciences - Helsinki 1961 p. 3 - 13.
- 6) G. FANT - Acoustic Theory of Speech Production Mouton & Co 1960.
- 7) G. ROSEN A Dynamic Analog Speech Synthesizer  
in J. Acoust. Soc. Am. 30 (1968) p. 20 - 209.

un intérêt certain pour les recherches phonétiques. Ces appareils, à leur tour, ont fait appel à l'ordinateur : la DAVO et la DANA furent contrôlées dès 1962 par un computer TX-0 (8). Les Laboratoires BELL ont même mis au point à une date plus récente un simulateur qui permet, d'entendre la parole artificielle et d'obtenir simultanément le mouvement de l'appareil phonatoire analogique sur l'écran d'un oscilloscope.

Avec l'introduction dans les sciences phonétiques de l'ordinateur sont nés enfin de nouveaux systèmes à génération de parole syn-hétique. Les uns, qu'on peut encore rattacher aux véritables méthodes de synthèse, en ce sens que les phonèmes sont bien recréés artificiellement, partent d'éléments « microscopiques » de la parole, c'est-à-dire d'un nombre restreint (600 éléments) de combinaisons de deux phonèmes (appelées selon le cas « diphones », « digrammes », ou « logatomes » préalablement mis en mémoire, pour reformer la voix (9). D'autres qui sont très intéressants sans doute par les applications dont ils sont susceptibles, mais qui ne peuvent plus être considérés comme des techniques de synthèse reconstituent les messages par assemblage d'éléments « macroscopiques » c'est-à-dire de mots pré-enregistrés (10).

Si l'on ne peut que se réjouir de ces orientations nouvelles que les techniques contemporaines proposent à la parole artificielle, il faut bien reconnaître toutefois qu'un certain nombre de difficultés subsistent encore.

Le reproche principal que l'on pourrait faire sans doute à la plupart des méthodes de synthèse, c'est que la préparation du program ne nécessite beaucoup trop de temps. La parole est un *acte spontané* : c'est cette spontanéité même qu'il faudrait pouvoir sauvegarder. Cela signifie qu'il faudra s'efforcer à l'avenir de substituer à la synthèse précédée d'une analyse une *synthèse immédiate*. En ce sens l'automate de FABER et le VODER qui pouvaient être commandés au moyen d'une série de touches ont satisfait sans doute le mieux ce rêve de faire parler une machine que l'humanité poursuit depuis des siècles. A. MOLES évoquant cette difficulté n'hésitait pas à qualifier l'idée de base dont procèdent ces appareils de « mythe dynamique ». C'est le rôle du philosophe ajoutait-il, de courir après les chimères...(Bibl. 20) Qu'elles partent d'éléments « macroscopiques » ou « microscopiques » les méthodes de synthèse ont aussi ceci en commun qu'elles opèrent un *découpage artificiel du continuum sonore*. La conséquence, c'est qu'elles n'arrivent pas toujours à ajuster les syllabes et à réaligner la continuité du discours. C'est le reproche que l'on faisait précisément au VODER..., l'inconvénient majeur des méthodes de synthèse immédiate... (les méthodes paramétriques donnant à cet égard des résultats bien meilleurs). Si ce découpage par lequel l'homme appré-

8) J. B. DENNIS - Computer Control of an Analog Vocal Tract - Proc. Stockholm Speech Comm. Seminar R.I.T. Stockholm, Sweden sept. 1962 cf. aussi F. S. COOPER - Instrumental Methods for Research in Phonetics in Proceedings of the Fifth International Congress of Phonetic Sciences Münster 1964 S. Karger - Basel p. 152.

9) cf. à ce sujet J. QUINIO et D. TEIL - La synthèse de la parole par ordinateur à partir de diagrammes phonétiques in Revue d'Acoustique 3e année n° 9 - 1970 pp. 28 - 32  
et J. S. LIENARD - Historique et Réalisations actuelles in Revue d'acoustique - 3e année - n° 11 - 1970 pp. 204 - 213.

10) cf. J. PONCIN

Synthèse de parole par mots in Revue d'Acoustique - 3e année n° 9 - 1970 pp. 17 - 20.

hender le monde pour le reconstruire n'a encore jamais été poussé aussi loin jusqu'à ce jour. Puisque nous raisonnons aujourd'hui en termes de « *Logatomes* » c'est-à-dire d'une *théorie atomiste du langage*, nous sommes en droit d'espérer toutefois que dans un avenir très proche, grâce aux mémoires des grands calculateurs, nous disposerons d'une science combinatoire capable d'intégrer assez rapidement ces atomes et de réaliser la continuité qui caractérise le message parlé. L'Institut de Linguistique et de Phonétique de l'Université d'Alger dispose déjà d'un pareil équipement pour la réalisation d'un Arabe Fondamental ; tout permet de penser qu'il pourra servir à l'avenir également à cette fin.

Le développement étonnant des méthodes de synthèse au cours des dernières années et l'intérêt considérable qu'on leur porte s'expliquent par le nombre même des applications pratiques dont les techniques de parole artificielle sont susceptibles. Elles paraissent aujourd'hui particulièrement prometteuses et touchent aussi aux domaines les plus divers. Il y a quelques temps déjà fut mis au point au C. N. E. T. un *Vocoder expérimental* pour la rééducation des sourds muets (Bibl. 10) c'est-à-dire un appareil reproduisant à la place des fréquences naturelles, des signaux dont les fréquences figurent dans la gamme audible du malade. Ces applications intéressent l'industrie au plus haut degré. Dans les colonnes publicitaires des récents numéros de la Revue d'Acoustique, Philips présente un *amplificateur à sons de compensation*. C'est-à-dire un appareil modifiant comme le précédent la structure acoustique de la parole en émettant des sons graves représentatifs des sons aigus inaudibles de l'enfant sourd. *L'audio-phonologie* n'est qu'un débouché entre bien d'autres. Les techniques de *télécommunications* tendent depuis longtemps à tirer parti de ces recherches qui devraient permettre à plus ou moins longue échéance la compression ultime des *bandes passantes* téléphoniques. Dans le cadre du projet Platon de central téléphonique électronique, le Centre National d'Etudes des Télécommunications de Lannion travaille dès à présent, à la mise en exploitation d'un système de parole synthétique par ordinateur (un calculateur doit desservir simultanément plusieurs synthétiseurs) pour le *service de la télétaxe* (Bibl. 102). Certaines applications présentent pour les constructeurs un enjeu particulièrement considérable. C'est le cas par exemple des techniques de reconnaissance de la parole qui dérivent des méthodes de synthèse et qui débouchent comme on sait sur la *commande automatique des machines* et sur la *traduction automatique*. DREYFUS-GRAF qui travaille déjà depuis un certain temps dans ce domaine présente dans un des derniers numéros de la Revue d'Acoustique (Bibl. 65) un appareil capable de reconnaître un certain nombre de mots : le *phonacteur* dont on nous dit qu'il sera suivi à une date ultérieure d'un *phonétographe* et même d'un *phonémographe*. Nous débouchons ainsi sur un autre « mythe dynamique » celui de la *dactylo automatique* dont personnellement nous regretterions l'apparition... Que de pareilles perspectives intéressent l'industrie, cela est évident... La plupart des grandes firmes de machines à calculer ont commercialisé dès à présent ce que l'on peut appeler des unités de réponse vocale (Bibl. 18). On aura certainement eu l'occasion aussi de lire comme nous cet article très intéressant sur le « téléphone des profondeurs » (11) et le *bathyphone* mis au point par la Compagnie I.B.M. pour restituer au signal vocal déformé par l'atmosphère d'hélium dans laquelle se trouvent les plongeurs des grands fonds son intelligibilité et pour le rendre compréhensible pour un auditeur situé dans l'air.

11) cf. Le Monde 11 nov. 1970 · Courrier des Sciences, p. 14.

Ainsi, le jour est proche sans doute, où la synthèse de la parole et les multiples applications qui en découlent feront partie de notre existence quotidienne. L'aboutissement de ce rêve tant caressé par l'humanité de faire parler une machine qui remonte comme nous avons vu jusqu'à l'Oracle de Delphes, et même au-delà, figure d'ailleurs déjà au nombre des extrapolations que l'on se plaît à faire dès aujourd'hui sur l'avenir.

Un article paru dans *Informatique* portait il y a quelque temps déjà (12) ce titre révélateur : « L'Oracle de Delphes ou la prospective de l'an 2 000 »...

### RESUME

Cette communication a pour but de faire connaître à tous ceux qui s'intéressent à l'acte de la parole le synthétiseur paramétrique EVA III actuellement en service au laboratoire de l'Institut de Linguistique et de Phonétique de l'Université d'Alger, et qui nous permet d'obtenir aujourd'hui, après plus de trois siècles d'efforts, une voix artificielle de très haute qualité.

Un aperçu historique où l'auteur retrace après d'autres, mais dans une optique essentiellement dynamique, l'évolution des machines parlantes surtout depuis l'éveil de la pensée scientifique au 18<sup>e</sup> siècle, permet de situer cette technique encore récente et d'apprécier le degré de perfection qui a été atteint.

L'exposé de la méthode est au centre même de l'article et largement illustré en annexe. Il est suivi d'un tableau plus indicatif qu'exhaustif des applications très nombreuses dont la méthode est susceptible, la plupart débouchant d'ailleurs sur des domaines très concrets : les techniques des télécommunications, l'automation ou la fabrication de matériel d'audiophonologie.

La conclusion ouvre quelques perspectives nouvelles sur l'avenir, l'évolution des méthodes de synthèse à l'ère de l'ordinateur et les applications pratiques qui entreront demain dans la vie courante.

12) cf. O1 *Informatique études*, décembre 1968, p. 16.

## الخلاصة

تركيب الكلام : تقرير لجلسة عمل عقدت حول موضوع الآلة الكهربائية المحاكية للصوت الطبيعي ( أيضا III )

### عرض منهاج

ان العرض من هذا العرض هو ان يعرف كل من يتم بدراسة حدوث الكلام . بآلة التركيب البارامترى ايما III الجاري استعمالها الآن بمعهد العلوم اللسانية والصوتية بجامعة الجزائر . وهي تمكنتنا اليوم من الحصول على صوت انساني مصطنع في غاية الجودة . ويأتي صاحب المقالة بلمحة تاريخية يتعرض فيها بدوره . لكن بكيفية أكثر دينامية لتطور الآلات الحديثة للكلام وذلك منذ أن استيقظ الفكر العلمي ( الأوربي ) في القرن الثامن عشر . وتمكنا هذه اللمحة من احلال صناعة التركيب الصوتي . الحديثة العهد . محلها من التطور العلمي كما تمكنا من ان نعرف مقدار ما بلغته هذه الصناعة من التقدم والاتقان . ومعظم هذا البحث يحتوي على عرض المنهاج في التركيب الصوتي وقصدت برسوم موضحة عديدة . ويتبعه جدول يشير بدون استقصاء الى التطبيقات الكثيرة جدا التي تترتب على استعمال هذا المنهاج اكثرها تفضي الى مجالات عملية متنوعة جدا . وذلك مثل تقنيات المواصلات وتسيير الآلات الذاتي وصنع الأدوات المستعملة في الطب السمعي الصوتي وغير ذلك .

ويختم المؤلف مقاله بالتشوف الى ما سيظروا من جديد في المستقبل في هذا الميدان ، والى تطور منهاج التركيب الصوتي في عهد الدماغ الإلكتروني ( الرتبة ) ، والتطبيقات المفيدة التي سوف يعود الناس على استعمالها في حياتهم اليومية .

## SUMMARY

### Speech Synthesis

Report on an experiment with the Electral Vocal Analog (EVA III)

First approach to a method

The aim of this paper is to introduce to all those interested in speech the parametric synthesizer EVA III, presently being used at the Institute for Linguistics and Phonetics of the University of Algiers ; this device makes it possible, after more than three centuries of research, to arrive at an artificial voice of very high quality.

Although the author is not the first one to present a short history of that research, he does it from a substantially dynamic point of view, showing us the evolution of speaking machines from the awaking of scientific thought in the 18 th century ; this makes it possible to situate the still recent technique of synthesizing, and to appreciate the degree of perfection which has been attained.

The explanation of the method of synthesis occupies the main part of this paper, and is substantially illustrated by additional graphs. At the end of the paper, we find an indicative rather than exhaustive table of the numerous applications of the method. Most of them lead to concrete areas : telecommunication techniques, automation, or the creation of audio-phonological material.

The conclusion opens some new prospects on the future, the evolution of synthesis methods in the new computer era, and eventually practical applications which will soon be part of everyday life.

## BIBLIOGRAPHIE

## OUVRAGES GENERAUX :

- 1) STRAKA Georges *Album Phonétique*  
Les Presses de l'Université de Laval (Québec 1965).
- 2) SIMON Péla *Les consonnes françaises*  
*Mouvements et positions articulatoires à*  
*La lumière de la radiocinématographie*  
Klincksieck - Paris, 1967
- 3) PEDERSEN Holger *The discovery of Language*  
trad. John Webster Spargo Indiana University Press Bloomington  
et London 1967.
- 4) MOLES A. et VALLANCIEN B. *Phonétique et phonation*  
Masson, Paris, 1966.
- 5) FLANAGAN James *Speech Analysis Synthesis and Perception*  
Springer Verlag - Berlin - Hèldelberg - New-York 1965.
- 6) DELATTRE Pierre *Studies in French and Comparative Phonetics*  
Mouton, 1966, London - The Hague - Paris.
- 7) FANT Gunnar *Acoustic Theory of Speech Production*  
Mouton et Cie Gravenhage, 1960.
- 8) MALMBERG Bertil *Manual of Phonetics*  
North-Holland Publishing C° - Amsterdam 1968.
- 9) LEHISTE Ilse *Readings in Acoustic Phonetics*  
The (M.I.T. Press 1967).
- 10) PIMONOW Léonid *Vibrations en Régime Transitoire*  
*Analyse Physique et Physiologique*  
Dunod, Paris, 1962.
- 11) DEWEZE A. *Traitement de l'Information Linguistique*  
Dunod, Paris, 1966.
- 12) DENES Peter *La chaîne de la communication verbale*  
(Lab. Bell Telephone).



13) ROUSSELOT Abbé P. *Principes de Phonétique expérimentale*  
Paris, 2<sup>e</sup> éd. 1924.

14) CHAPUIS A. *Les Automates*  
Neuchâtel, 1949.

#### HISTORIQUE :

15) METTAS Odette Aperçu historique sur les appareils de synthèse de la parole (texte d'une conférence faite à la Société de Cybernétique à Paris 22/5/65)  
in *Travaux de Linguistique et de littérature*  
Centre de Philologie et de Littératures Romanes de l'Université de Strasbourg III, 1 (1965) pp. 185-194.

16) FLANAGAN J. ouvrage cité pp. 166-291.

17) COOPPER Franklin *Speech Synthesizers*  
in *Proceedings of the 4th. International Congress of Phonetic Sciences*  
Helsinki, 1961.  
(Mouton et C<sup>o</sup> - The Hague - pp. 3-13)  
et *Instrumental Methods for Research in Phonetics*  
in *Proceedings of the 5th. International Congress of Phonetic Sciences*  
Münster, 1964 (S. Karger Basel, New-York) pp. 142-171.

18) LIENARD J. S. *La synthèse de la parole*  
*Historique et réalisations actuelles*  
in *Revue d'Acoustique*  
3<sup>e</sup> année n<sup>o</sup> 11 - 1970 pp. 204-213 (Cet article qui est excellent par sa présentation nous est parvenu alors que la présente étude était déjà à la correction. Nous avons essayé dans toute la mesure du possible de tenir compte des informations nouvelles qu'il apportait sur ce sujet).

19) FANT G. *Analysis and synthesis of Speech processes*  
in *Manual of Phonetics*  
Malmberg B. Edit. ouv. cité pp. 173-277.

20) MOLES A. et  
VALLANCIEN B. *Les méthodes de la Phonétique Expérimentale*  
ouv. cité pp. 15-60.

#### MACHINES PARLANTES DU 18<sup>e</sup> SIECLE :

21) HEMADINQUER J. J. *Machines parlantes peu connues du 18<sup>e</sup> siècle*  
in *Revue du son*  
nov. 1961.

- 22) KRATZENSTEIN Christian G. Tentamen Coronatum de Voce  
*Acta. Acad. Petrog.* 1780  
et trad. sur la naissance et la formation des voyelles  
*J. Phys.* 21, 358-380 (1782)
- 23) KEMPELEN W. Von *Le mécanisme de la parole suivi d'une description d'une machine parlante*  
Vienne, J. V. Degen, 1791.
- 24) DUDLEY H., TARNOCZY Speaking machine of Wolfgang von Kempelen, *J.A.S.A.* 22/2/50.

19<sup>e</sup> SIECLE :

- 25) WHEATSTONE Sir Ch The Scientific papers of Sir Charles Wheatstone  
*Physical Society of London*  
Tylor & Francis, 1879.
- 26) GABRIEL M. Machine parlante de M. FABER  
*Journal de Physique théorique et appliquée* 8, 1879.
- 27) HELMHOLTZ *Sensations of tone*, 1875.
- 28) KOENIG R. *Quelques expériences d'Acoustique*, Paris, 1882.
- 29) FREECE W. H., STROH A. On the synthetic examination of vowel sounds.  
*Proc. Royal Soc. London*, 28, 1879.

20<sup>e</sup> SIECLE :

- 30) MILLER *Science of Musical Sounds*.  
New-York : Macmillan, 1916.
- 31) STUMPF O. *Die Sprachlaute* - Berlin-Springer, 1929.
- 32) PAGET Sir Richard *Human Speech* - London, New-York : Harcourt, 1930.
- 33) RIESZ d'après Flanagan ouv. cité pp. 168-170.  
*Mechanical talker* (non publié, 1937).
- 34) STEWART J. Q. An Electrical Analog of the Vocal Organs  
*Nature* 110, 311-312 (1922).
- 35) FLETCHER H. Demonstration lecture introducing the new tone synthesizer.  
*Amer. J. Phys.* 14 - 1946.

- 36) WAGNER K. W. Ein neues elektrisches Sprechgerät zur Nachbildung der menschlichen Vokale.  
*Preuss Akad. Wiss. Berlin Abb.* 2, 44 p., (1936).
- 37) DUDLEY H. W. — Synthèse de la parole  
*Ann. P.T.T. Fr.* 27 (1938) pp. 136-142.  
trad. de BELL Lab. Rec. USA, 15, 4 (déc. 1936).  
— Remarking Speech *J.A.S.A.* 11 (1939) pp. 169-177.
- 38) DUDELEY H., RIESZ R. R., WATKINS S. A. S. *A synthetic speaker* *J. Franklin Inst., USA*, 227 (1939) pp. 739-764.
- 39) DUDLEY H. W. — The carrier nature of speech.  
*Bell syst. Tech. J.* 19 (1940) pp. 495-515.  
The Vocoder - electrical recreation of speech  
*J. soc. Mot. Pict. Engrs. U.S.A.*, 34 (1940) p. 272  
— The Vocoder remakes speech  
*Proc. Inst. Radio Engrs.* 28 (1940) pp. 1-47.
- 40) KOENIG W., DUNN H. K., LACEY L. Y. The Sound Spectrograph  
*J.A.S.A.*, 18 - pp. 19-49, 1946.
- 41) POTTER R. K., KOPP G. A., GREEN H. C. *Visible Speech*  
New-York : D. von Nostrand Co - 1947.
- 42) SCHOTT L. O. A Playback for Visible Speech  
*Bell Lab. Record.* pp. 333-339 (1948).
- 43) BORST J. M., COOPER F. S. Speech Research Devices Based on a Channel Vocodor  
*J.A.S.A.*, 29, (1957).
- 44) et BORST J. M. The Use of Spectrograms for Speech Analysis and Synthesis  
*J. Audio Eng. Soc.* 4, (1956) pp. 14-23.
- 45) LAWRENCE W. *The Synthesis of Speech from Signals which have a Low Information Rate*  
Comm. Theory éd. Jackson (London 1953).
- 46) FANT G. Modern Instruments and Methods for Acoustic Studies of Speech  
in *Proc. of the VIII international Congress of Linguistics*  
Oslo, 1958 pp. 282-358.  
et Recent Progress in Formant Synthesis of Connected Speech  
*J.A.S.A.*, 33, (1966) pp. 834-835.

## SYNTHETISEUR EVA III :

- 47) Documentation : MELPAR inc. Falls Church. Virginia 22046.
- INTERET DES METHODES DE  
SYNTHESE - APPLICATIONS -  
PHONETIQUE ACOUSTIQUE :
- 48) DELATTRE P. Les indices acoustiques de la parole  
in *Studies in French and Comparative phonetics*  
ouv. cité pp. 248-275.  
Le Jeu des transitions de Formants et la perception des conson-  
nes, idem pp. 278-286.
- 49) DELATTRE P., Acoustic Loci and Transitional Cues for Consonants  
LIBERMANN A. et  
COOPER J.A.S.A., vol. 27, (1955) pp. 769-793
- 50) DELATTRE P., (Haskins Laboratories)  
LIBERMANN A., Tempo of frequency change as a cue for distinguishing classes  
GERSTMAN L. et of Speech Sounds.  
COOPER F. *Journal of experimental Psychology*, vol. 52, (1956) pp. 127-137.
- 51) JANOS MARTONY, On the Analysis and Synthesis of Vowels and Fricatives  
CERDERLUND C., in *Proceedings of the 4 th. International Congress of Phonetic  
LILJENCRANTS J., Sciences*  
LINDBLOM B. ouv. cité pp. 208-214.
- 52) SHEARME et HOLMES An Experimental Study of the Classification of Sounds in Con-  
tinuous speech according to their Distribution in the Formant.  
1 - Formant 2 Plane.  
in *Proceedings of the 4 th. International Congress of Phonetic  
Sciences*  
ouv. cité pp. 234-241.
- 53) DELATTRE P., An experimental study of acoustic determinants of vowel color ;  
LIBERMANN A., observations on one - and two - formant Vowels synthesized  
COOPER F. et from spectrographic patterns. *Word* 8 : 195-211 (1952).  
GERSTMAN L.
- 54) MEYER-EPPLER W. Zum Erzeugungsmechanismus der Geräuschlaute  
*Z. Phonetik u. allgem. Sprachwiss.*  
Dtsch., 7 (1953) pp. 196-212.
- 55) HECKER Studies of Nasal Consonants with an  
Articulatory Speech Synthesizer  
*J.A.S.A.*, 34, (1962) pp. 179-188.

- 46) HOUSE Arthur et FAIRBANKS The Influence of Consonant Environnement upon the secondary Acoustical Characteristics of vowels  
*J.A.S.A.*, vol. 25, (1953- pp. 105-113
- 57) HARRIS K. S., HOFFMAN H. S., LIBERMANN A. M., DELATTRE P., COOPER F. S. Effect of the third - formant transitions on the perception of the voiced stop consonants.  
*J.A.S.A.* 30, (1958) p. 122.
- 58) WARNS O Die Verständlichkeit synthetischer Sprache  
*Frequenz*, Dtsch., 11 (1957) pp. 109-175.
- 59) TARNOCZY T. Über das Individuelle Sprachspektrum in *Proceedings of the 4 th. International Congress of Phonetic Sciences* ;  
ouv. cité pp. 259-265.

## APPLICATIONS :

- 60) FLANAGAN J. L. Bandwith and Channel capacity necessary to transient the formant information of speech  
*J.A.S.A.*, 28, (1956) p. 592.
- 61) ENDRES W. Methoden der Sprachanalyse und Sprachsynthese und ihre Bedeutung für die Nachrichtentechnik.  
*Fermeld. Ing.*  
11 (nov. 1957).
- 62) VILBIG F., MAASE K. H. Über einige Systeme zu Sprachbandkompression  
*Nach. tech.*  
Fachber. 3 (1956-57).  
et Some systems for speech-band compression  
*J.A.S.A.*, 28, (1956) p. 573.
- 63) HALSEY R. J., SWAFFIELD J. Analysis - synthesis - telephony, with special reference to the Vocoder.  
*J. Inst. Elec. Engrs G. B.*,  
II, 95 (1948) pp. 391-411.
- 64) BELEVITCH V. *Langage des machines et langage humain*  
éd. Hermann, Paris, 1956.
- 65) DREYFUS-GRAF J. Phonétographe, Présent et Futur.  
*Bulletin technique P.T.T.*,  
1961, pp. 160-172.

- 66) OLSON H. F.,  
BELAR H.                    Printout System for the Automatic Recording of the spectral  
Analysis of Spoken Syllables  
*J.A.S.A.*, 62, (pp. 166-171).
- 67) CAMPANELLA S. P.        A Survey of Speech Bandwith Compression Techniques  
*IRE Trans.* AU-6 (sept. - oct. 1958) pp. 104-116 (23 réf. bibl.)
- 68) SIMMONS P. L.            Automation of Speech, Speech Synthesis and synthetic Speech  
A Bibliographical Survey from 1950-60  
*I.R.E. Trans.* AU-9 (nov.-déc. 1961) pp. 191-196
- 69) HOWELL A. S.,  
SCHNEIDER G. O. K.,  
STUMP T. M.                A Military Semi-Vocoder for Analog Telephone Transmission  
*I.A.S.A.* 33, (1963) (A) 1961.

## PERCEPTION ET INTELLIGIBILITE :

- 70) DELATTRE P.             De la hiérarchie des indices acoustiques pour la perception de la  
parole  
in *Proceedings of the 5 th. International Congress of Phonetic  
Sciences*,  
ouv. cité pp. 244-251.
- 71) FOURQUET J.            Perception et reproduction des traits pertinents  
in *Proceedings of the 4 th. International Congress of Phonetic  
Sciences*  
ouv. cité pp. 292-295.
- 72) COHEN A. et  
KATWIJK A. F. J. von        The Perception of Phonemes as a Function of Acoustic and  
Distributional Cues  
in *Proceedings the 5 th. International Congress of Phonetic  
Sciences*  
ouv. cité pp. 598-610.
- 73) WINCKEL F. von        Perzeptive Grenzen der Phonem - Unterscheidung  
in *Proceedings of the 5 th. International Congress of Phonetic  
Sciences*  
ouv. cité pp. 582-588.
- 74) LIBERMANN A.,  
DELATTRE P.,  
COOPER F. S.                — The role of selected stimulus - variables  
in the perceptoin of the unvoiced stop consonants  
*Amer. J Psychol.* 65 pp. 497-517 (1952).  
— Acoustic loci and transitional cues for consonants  
*J.A.S.A.*, 27, pp. 769-774, (1955).

- 75) LIBERMANN A.,  
DELATTRE P.,  
COOPER F.,  
GERSTMAN L. The role of consonant - vowel transitions  
in the perception of the stop and nasal consonants  
*Psychol. monogr.* 379 : 1-4 (1954).
- 76) HOFFMAN H. S. Study of some cues in the perception of voiced stop consonant  
*J.A.S.A.* 30, 11 : 1035-1041 (1958).
- 77) COOPER F. S.,  
DELATTRE P.,  
LIBERMANN A. M.,  
BORST J. M.,  
GERSTMAN L. J., Some experiments on the perception on synthetic Speech sounds  
*J.A.S.A.*, 24, (1952) pp. 597-606.
- 73) FAURE G.,  
INGEMANN F., LISKER L., DELATTRE P.,  
COOPER F. L'intonation et l'identification des mots dans la chaîne parlée  
in *Proceedings of the 4<sup>th</sup>. International Congress of Phonetic  
Sciences*  
ouv. cité pp. 598-610.
- 79) LIBERMANN A. M.,  
INGEMANN F.,  
LISKER L.,  
DELATTRE P.,  
COOPER F. Minimal Rules for Synthesizing Speech  
in *J.A.S.A.* vol. 31 (1959) pp. 1490-1499.

## APPLICATIONS :

- 80) LAFON J. C.,  
CHEVALLIER J., Prothèse et phonétique  
*Philips audiométric*, Fr., 13 (1960)
- 81) MILLER R. J. Audiotory test with synthetic vowels  
*J.A.S.A.*, 25, (1953) pp. 114-121.
- 82) PIMONOW L. Quelques aspects théoriques sur la relation entre les troubles  
auditifs et les informations.  
*in Vibrations en régime transitoire* ;  
ouv. cité pp. 275-280.
- 83) SCHROEDER M. R. et A Vocoder for transmitting 10 kHz speech over a 3, 5 k Hz  
DAVID J. Channel in *Acustica, Suisse*, 10 (1960) pp. 35-43.
- 84) FLETCHER H., GALT Perception of Speech and its relation to telephony  
R. *J.A.S.A.*, 22, (1950) p. 89.

ETUDE DES FAITS D'ACCENT ET DE  
MELODIE :

- 85) RIGAULT A.            Rôle de la fréquence, de l'intensité et de la durée vocaliques dans la perception de l'accent en français.  
in *Proceedings of the 4 th. International Congress of Phonetic Sciences*  
ouv. cité pp. 735-749.  
et Marges licites et seuil critique des variations de l'intonation dans la phrase déclarative en français in *Proceedings of the 5th. International Congress Phonetic Sciences*  
Münster, 1964.
- 86) METTAS Odette        Les facteurs ecto-sémantiques du discours et leur caractérisation par synthèse.  
in *Phonétique et pbonation* pp. 177-188.  
et Etude sur les facteurs ecto-sémantiques de l'intonation en français  
in *Travaux de Linguistique et de Littérature*  
Centre de Philologie et de Littératures romanes de l'Université de Strasbourg (1963) - Klincksieck éd. pp. 143-149.
- 87) ULDALL E.            Ambiguity ; Question or statement ?  
" Are you asking me or telling me ? "  
in *Proceedings of the 4 th. International Congress of Phonetic Sciences*  
pp. 779-783.
- 8) FONAGY                Zur Gliederung der Satzmelodie  
in *Proceedings of the 5 th. International Congress of Phonetic Sciences*  
pp. 281-287.
- 89) PIKE K. L.            On the Grammar of Intonation  
in *Proceedings of the 4 th. International Congress of Phonetic Sciences*  
pp. 105-117.

ORIENTATIONS NOUVELLES :  
ORDINATEURS ET VOCODERS.

- 90) FLANAGAN J. L.,        Computer Simulation of a Formant Vocoder Synthesizer  
COKER CH., BIRD        in *J.A.S.A.*, 35, 2003 (A) 1962.
- 91) KELLY J. L.            An Artificial Talker driven from a Phonetic Input.  
GERSTMAN L. J.        *J.A.S.A.*, 33, 835 (A) 1961 (avec IBM 7090).



- 92) KELLY J. L., GERSTMAN L. J. et LOCHBAUM C. Speech synthesis - *Proc. Stockholm Speech Comm. Seminar*, R.I.T. Stock., Sweden, sept. 1962
- 93) KELLY J. L., GERSTMAN L. J., VYSSOTSKY V. A. A Block Diagram Compiler *BELL system. Tech.* 40 pp. 669-676 (1961).
- 94) GOLDEN R. M. Digital Computer Simulation of a Sampled - Data Voice Excited Vocoder - *J.A.S.A.*, 35, 1358-1366 (1963).
- 95) HOLMES et SHEARME Speech Synthesis by rule controlled by a small low-speed digital computer *J.A.S.A.*, 35 - 1911 (A) 1963.

ORDINATEURS ET SYNTHETISEURS  
ANALOGIQUES :

- 96) STEVENS, KASOWSKI, FANT An Electrical Analog of the Vocal Tract *J.A.S.A.*, 27, (1953) pp. 734-742.
- 97) DUNN H. K. The Calculation of Vowel Resonances and an Electrical Vocal Tract - *J.A.S.A.*, 22, (1950) pp. 740-753.
- 98) ROSEN G. A Dynamic Analog Speech Synthesizer in *J.A.S.A.* 30 (1968) pp. 20-209
- 99) DENNIS J. B. Computer Control of an Analog Vocal Tract *Proc. Stockholm Speech Comm. Seminar* R.I.T. Stockholm, Sweden sept. 1962

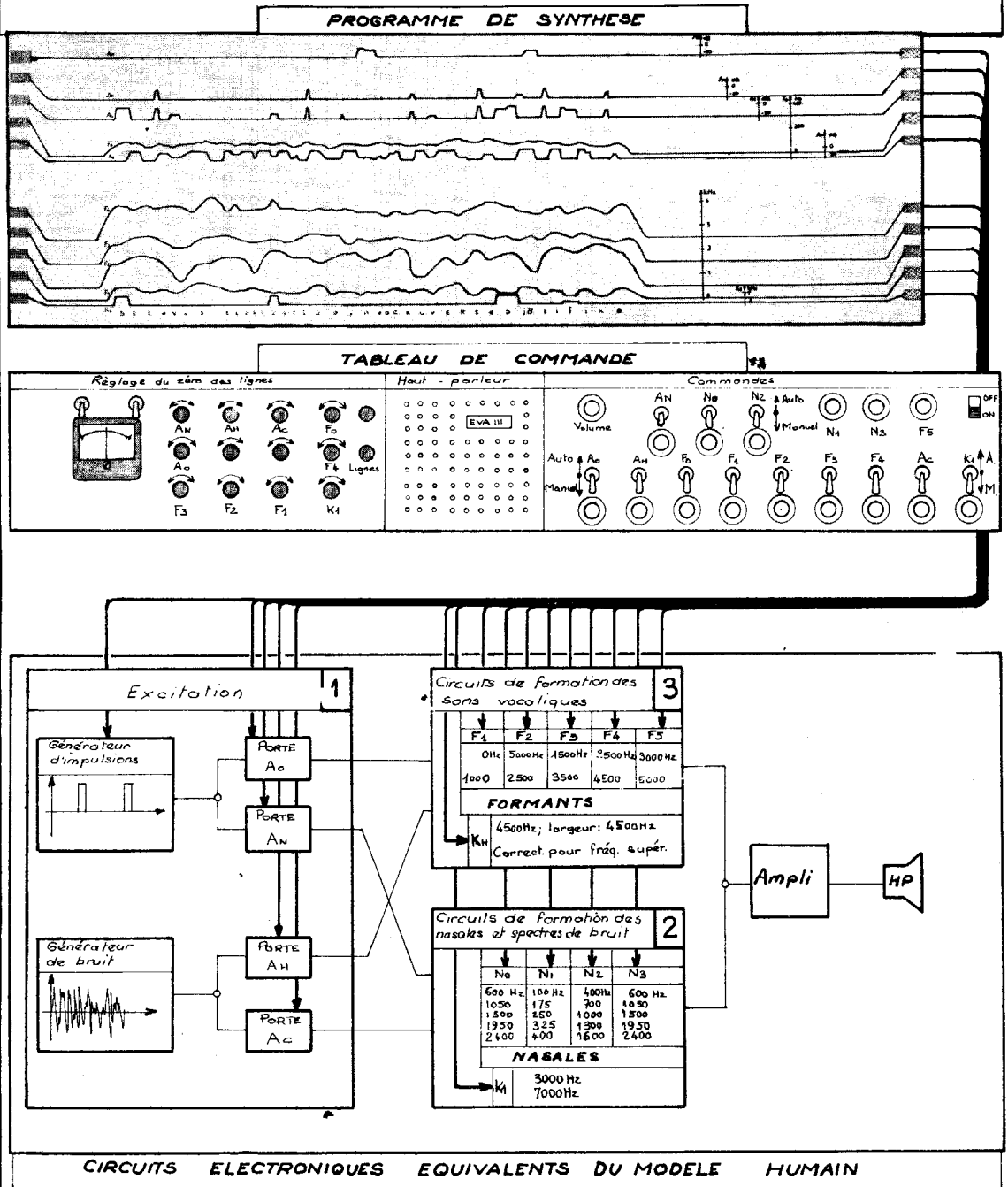
SYNTHESE PAR ORDINATEURS :

- 100) MOLES A. Les Méthodes de la phonétique expérimentale in *Phonétique et phonation* pp. 50-55
- 101) QUINIO et TEIL La synthèse de la parole par ordinateur à partir de diagrammes phonétiques in *Revue d'Acoustique* 3<sup>e</sup> année n° 9 - 1970 pp. 28-32
- 102) PONCIN J. Synthèse de parole par mots in *Revue d'Acoustique* 3<sup>e</sup> année n° 9 - 1970 pp. 17-20.
- 103) ESTES, MAXEY et WALKER Speech Synthesis from stored data *IBM Journal* 8, 2-12, 1964
- 104) COOPER F. S. Speech from Stored data *IEEE - Convention Record* 7, 137-149 (1963)

- 105) DENES P.                    The use of Computers for Research in Phonetics  
in *Proceedings of the 4 th. Intern. Congress of Phonetic Sciences*  
pp. 149-154.
- 106) MATHEWS M. V.            The effective use of digital simulation for speech process.  
*Proc. Seminar on Speech Compression and Proc.* vol.1  
(Air Force Cambridge Research Center, Mass. 1959).

# LE SYNTHETISEUR EVA III

## SCHEMA FONCTIONNEL

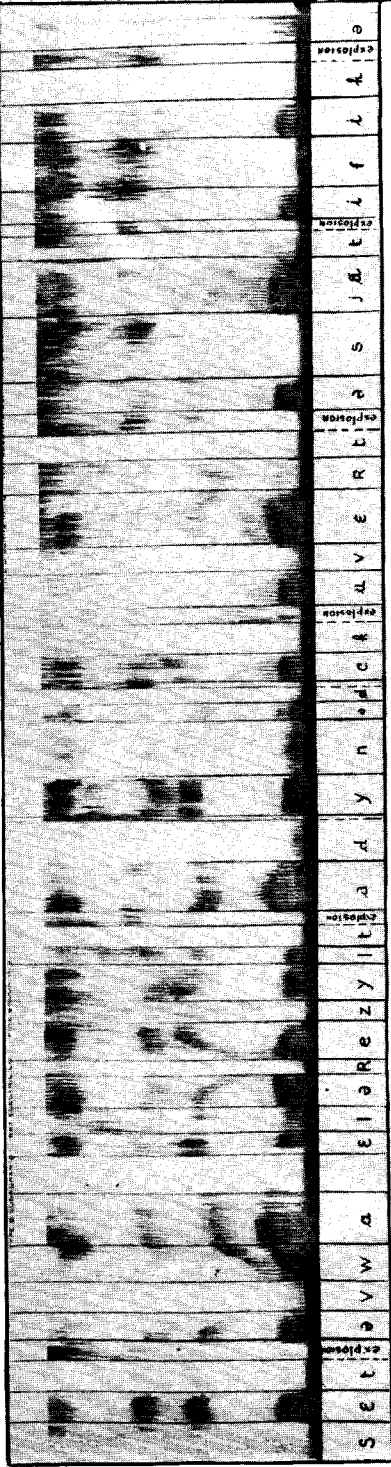




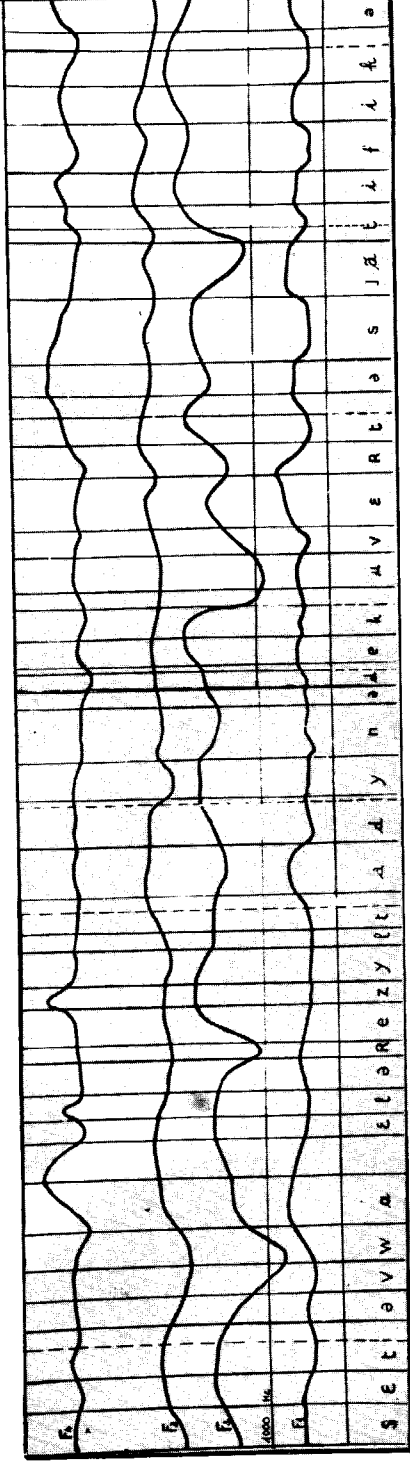
L'ELECTRICAL VOCAL ANALOG EVA III  
 REALISATION D'UN PROGRAMME DE SYNTHÈSE

P13

SEQUENCE: "Cette voix est le résultat d'une découverte scientifique".



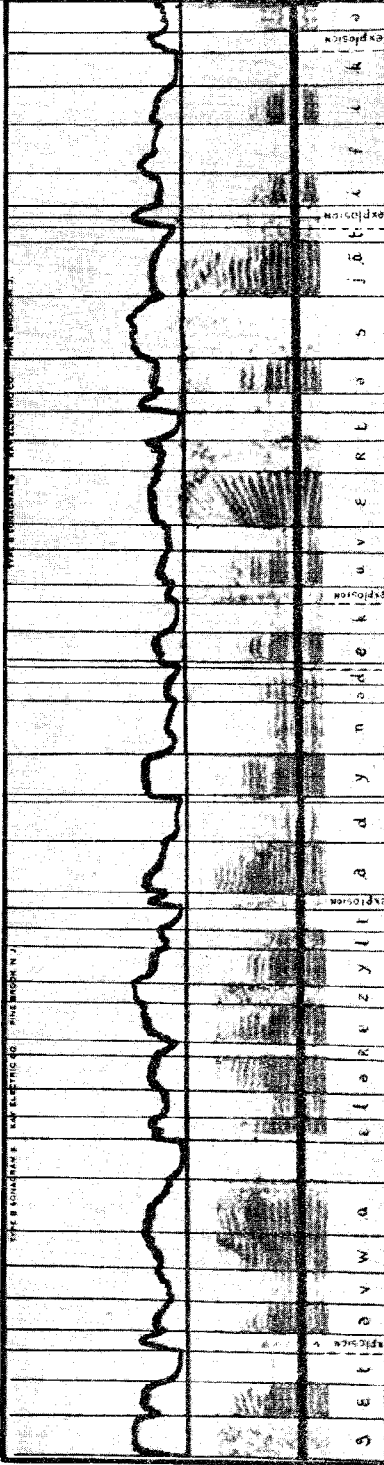
Sonagramme (Filtrage large: 300 cps - Echelle dilatée: 85-4000 Hz)



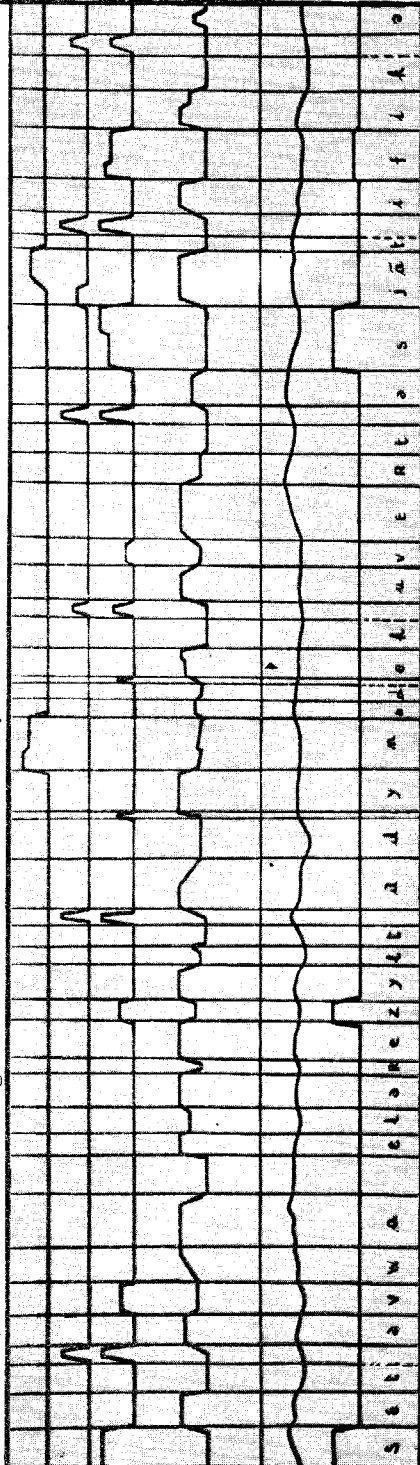
Graphes obtenus par décalque sur une feuille de plastique des lignes d'évolution des quatre premiers formants ( F1, F2, F3, F4 )

L'ELECTRICAL VOCAL ANALOG EVA III  
 REALISATION D'UN PROGRAMME DE SYNTHESE.

SEQUENCE: " Cette voix est le résultat d'une découverte scientifique " .



Sonagramme (Filtrage étroit: 45cps -Echelle dilatée: 85-4000 Hz)  
 pour le décalque de la ligne d'évolution du 5ème harmonique et l'obtention de la ligne mélodique (Fo)  
 Partie supérieure du tracé: ligne d'évolution de l'intensité.

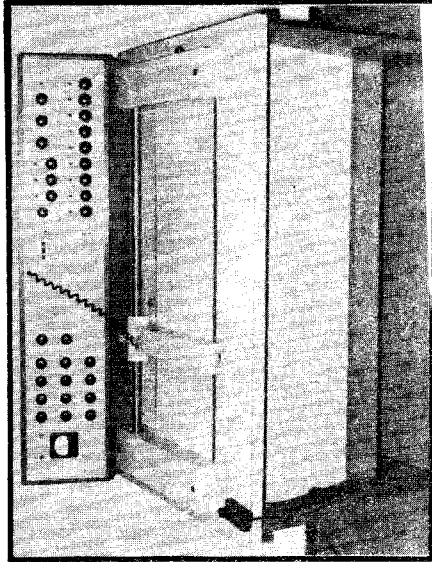


Graphes des paramètres K1, Fo, Ao, Ac, Ah, An .  
 Fo est obtenu par décalque (5ème harmonique) ; Ao, Ac, Ah, An s'obtiennent à partir de la ligne d'intensité;  
 K1 est construit en fonction des données de l'analyse.

L 'ELECTRICAL VOCAL ANALOG EVA III  
REALISATION D'UN PROGRAMME DE SYNTHESE

PI 5

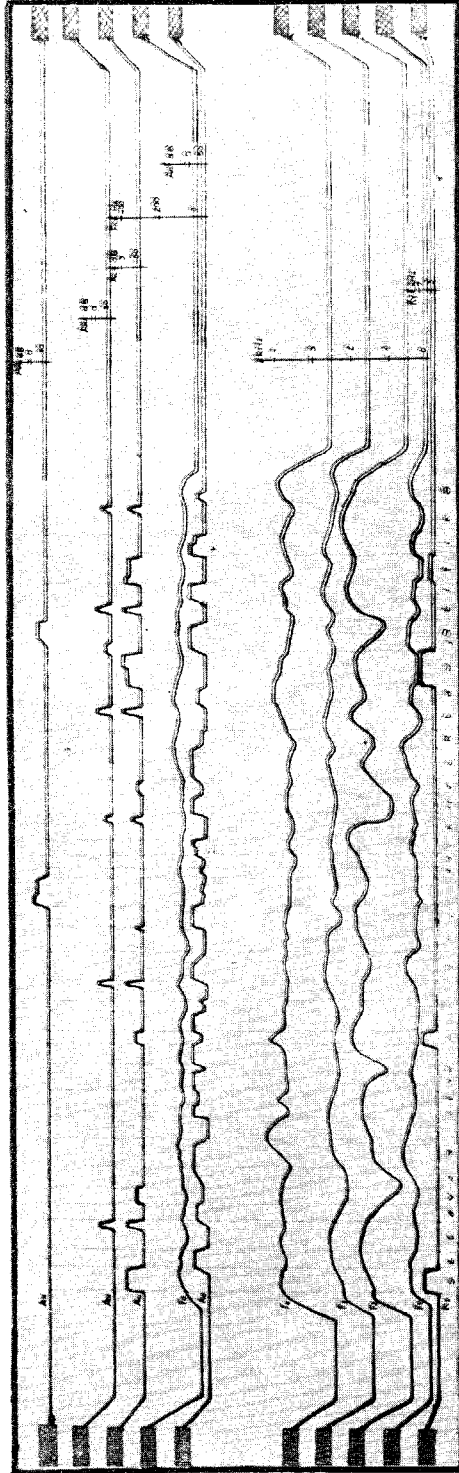
L'APPAREIL:



Le PROGRAMME comporte les dix PARAMETRES suivants:

- K1 : fréquences pour la détermination des spectres de bruits
- F1, F2, F3, F4 : lignes d'évolution des quatre premiers Formants (pour le système de filtres vocaniques)
- Ao : amplitude de la voix
- Fo : variations de la fréquence du Fondamental (ligne mélodique)
- Ac : amplitude des bruits des fricatives dont les bandes de fréquences sont déterminées par K1
- Ah : amplitude de bruits pour le système de filtres vocaniques (aspiration)
- An : amplitude des composantes nasales

LE PROGRAMME:

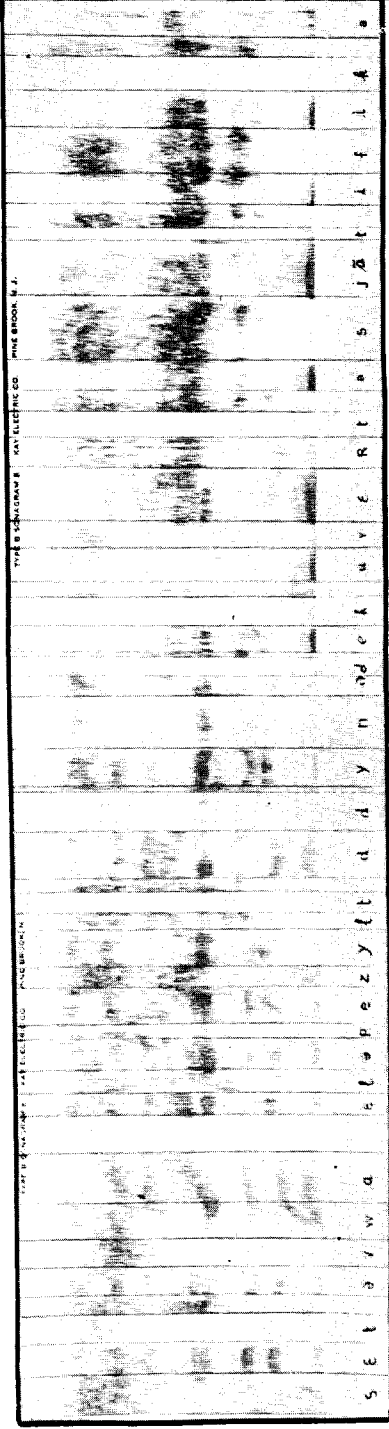


"CETTE VOIX EST LE RESULTAT D'UNE DECOUVERTE SCIENTIFIQUE"

L'ELECTRICAL VOCAL ANALOG EVA III  
REALISATION D'UN PROGRAMME DE SYNTHESE

P/6

SEQUENCE : "Cette voix est le résultat d'une découverte scientifique!"



ORIGINAL : Sonogramme en filtrage large : 300 cps - Echelle : 85 - 8 000 Hz



PAROLE SYNTHETIQUE : Sonogramme en filtrage large : 300 cps - Echelle: 85 - 8000 Hz