

E. LEIPP

INFORMATION SEMANTIQUE ET PAROLE

ESSAI D'UNE GESTALTTHEORIE



JUIN 1966

N°22

G A M

BULLETIN DU GROUPE D'ACOUSTIQUE MUSICALE

Faculté des Sciences — 8 rue Cuvier — PARIS 5^e

1°) REUNION DU 28 JUIN 1966

Etaients présents :

M. le Professeur SIESTRUNCK, Président
M. LEIPP Secrétaire général; Mlle CASTELLENGO, Secrétaire.

puis, par ordre d'arrivée,

Mme LEIPP; Mlle M. LEIPP (Etudiante lettres); M.J.S. LIENARD (Ingénieur A. et M.); M. FRANCOIS (E.D.F.); M. GIRVES (O.N.E.R.A.); M. DE MARIS (O.N.E.R.A.); Mme MOTTIER (psychologue); M. CANY (Psychologue); M. PIVA (Etudiant en médecine); M. CAMION (Architecte); M. FORET (Ex-chef de la Garde); M. BOUILLON (Professeur Sourds Muets); Dr VERNANT (Dr en médecine); Mme ROTTERDAM (Médecin); M. CHULLIAT (Professeur Sourds Muets); Mme METTAS (Phonéticienne); Mme CHARNASSE (CNRS); Mme et M. DUPARCOQ (Revue Musicale); Dr LANGINIER (Orthophoniste); M. MOUTET (ONERA); M. SAPALY (Labo. Méca. Phys. Fac. Sciences); M. A. MEYER (Licencié de physique); Mme BOREL MAISONNY (Orthophoniste); Melle DINVILLE (Orthophoniste); M. MARCIE (Dr en médecine); M. J. DUBOIS (Prof. Fac. Lettres HAUTERRE); Dr. HECAEN (Dr en médecine); Melle REAUDIE; Melle PRADEL (Conservatoire de Musique); M. SADEK; M. BATTISSIER (SIERE); M. J. DESCHAMPS; M. METAIS (Professeur); M. CARCHEREUX (Maitre luthier); Melle Edith WEBER et M. STOQUER (Institut de musicologie); M. (Kings College); M. AKIRA TAMBA (compositeur); Dr J. BARBIZET (Dr en médecine); M. HIRSCHMANN (Psychologue); M. M. GAUTHIER (CNRS); M. BLONDELET (Dir. Ets BUFFET CRAMPON); Mme NYEKI (Phonothèque Nationale); Mme Yvette GRIMAUD (CNRS); Dr. DORGEUILLE (Dr en médecine); Dr KADRI (phoniâtre); M. BOLTANSKI (Institut de phonétique); Mme COUTURE (Professeur Fac.Sc. Paris); M. POUBLANC (Médecin biologiste); M. DUBUC (Musée des Arts et Traditions populaires); Mme BIZAGUET et Dr VEIT (Laboratoire de correction auditive); Melle GRISON; Dr J. PERROT (Dr ès lettres); Mme M.J. CHAUVIN (Courrier Musical de France);

Excusés : M. GAUTHIER (Vice Doyen Fac. Sciences Paris); M. J. CHAILLEY (Dir. Institut Musicologie); Mme de CHAMBURE; Dr APAP; M. FAYEULLE; M. LEON; M. GRONIER; M. GUEN; M. BASCHET, M. PUJOLLE; M. DEWEVRE; Dr VALLANCIEN; M. GILOTAUX; M. SELMER; M. CANAC; Mme de BOISSIEU; Mme FULIN; M. CAPELLE; M. ACOULON; M. BERNARD; M. HOCQUENGHEM; Melle S. CORBIN; M. MAILLOT.

2°) Nous avons appris la mort de HERMANN SCHERCHEN, chef d'orchestre et compositeur. Dans le domaine de l'acoustique il avait créé le Studio expérimental de GRAVESANO (Suisse) et dirigeait une revue d'acoustique musicale, les GRAVESANER BLÄTTER, unique en son genre.

3°) Nous avons eu le plaisir d'accueillir au laboratoire, en Août, M. YOSHIKAZU TAKEDA, de la NIPPON GAKKI (YAMAHA) accompagné de M. BLONDELET, et de M. le Prof. William R. BENNET de l'Université de Columbia accompagné de son fils.

...../

P L A N

I	- HISTORIQUE	p. 1
II	- LE SYSTEME PHONATOIRE ET LA STRUCTURE PHYSIQUE DE LA PAROLE.	p. 5
	1°) Le système phonatoire	p. 5
	2°) La structure physique des signaux phonétiques	p. 7
III	- LA NATURE INFORMATIVE DU MESSAGE PARLE	p. 9
IV	- L'EXTRACTION DE L'INFORMATION SEMANTIQUE ET SON INTEGRATION PAR L'HOMME	
	1°) L'expérience de la guimbarde	p. 10
	2°) Le rôle du spectre de raies	p. 12
	3°) La voix chuchotée	p. 13
V	- GESTALTTHEORIE ET PAROLE	p. 16
VI	- ELEMENTS PHONETIQUES ET SYNTHÈSE DE LA PAROLE	p. 16
VII	- CONCLUSIONS	p. 22

ESSAI D'UNE GESTALTTHEORIE

Résumé de l'exposé fait au GAM le 28 Juin 1966

par M. E. LEIPP

Cet exposé résume les résultats que nous avons obtenus au Laboratoire d'acoustique dans le domaine de la parole. Le groupe de recherche est composé de :

M. LEIPP, Melle CASTELLENGO, M. J.S. LIENARD.

L'appareil de synthèse (ICOPHONE) réalisé à l'occasion de cette étude a été construit et mis au point par la Société DUFOUR ELECTRONIC en collaboration avec les techniciens du Laboratoire de Mécanique Physique de Saint-Cyr-l'Ecole.

I. HISTORIQUE

Les problèmes posés par la voix, la parole et le langage articulé ont suscité la curiosité et l'intérêt de nombreux chercheurs depuis l'antiquité. Nous en avons des preuves certaines depuis ARISTOTELE (4^e siècle avant notre ère)... Après lui, GALIEN s'intéresse également au mécanisme de production de la voix et nous voyons apparaître de bonne heure des tentatives de simulation de la parole.

Plus près de nous, MERSENNE se passionne pour ces problèmes et consacre tout le chapitre I de son HARMONIE UNIVERSELLE à l'étude de " de la voix, des parties qui servent à la former, de sa définition, de ses propriétés et de l'ouïe ".

La lecture de ce texte est très instructive et nous renseigne sur l'état de la question au 17^e siècle. Selon son habitude.

l'auteur, esprit curieux, pose de nombreuses questions fort pertinentes. Ce sont, tout au long, des : " expliquer pourquoi expliquer par quels A savoir si....

Si les réponses ne sont pas toujours correctes, c'est parce qu'on ne disposait alors que d'informations très sommaires et d'appareillages acoustiques rudimentaires. Mais il n'est pas douteux que l'auteur n'ait longuement réfléchi au problème.

Dans le livre sixième du même ouvrage, qui se rapporte à l'orgue, il revient sur cette question et nous propose des solutions pour la synthèse de la parole. Le titre de la XXXVI^e proposition est en effet :

" Expliquer comme il faut construire les jeux d'orgue pour prononcer les voyelles, les consonnes, les syllabes et les diction. " Le texte mérite d'être lu en entier; nous en retiendrons quelques passages :

...." Il est premièrement certain qu'il n'y a que les tuyaux à anches qui puissent servir à la prononciation des voyelles, des consonnes et des syllabes, comme l'enseigne l'expérience..."

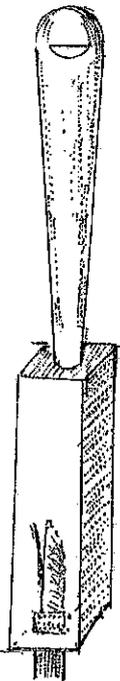
.... Mais il faut remarquer que la configuration et la modification du son se fait principalement par les différentes figures du corps du tuyau, et non par celle des languettes, ou des échallottes qui doivent toujours être de même façon, quoi que l'on puisse faire la languette plus épaisse ou plus déliée. Or le corps du tuyau peut être altéré en plusieurs manières "

MERSLANNE semble très bien avoir compris certains problèmes actuels (question des transitions phonétiques, théorie du locus etc...) et ses observations confirment souvent des aspects du problème qui nous apparaissent à l'observation des sonagrammes. Ainsi en est-il de certaines consonnes, les explosives par exemple, dont il dit :

" il semble que les consonnes ne sont autre chose qu'une certaine circonstance des voyelles " etc...

En fait, nous avons observé, à l'occasion d'analyses faites au laboratoire sur des jeux d'orgues de types très particuliers que les facteurs d'orgue savaient parfaitement réaliser, de façon empirique bien sûr, mais tout à fait systématique, des associations de formants simulant les voyelles. Ainsi nous avons remarqué sur un orgue CALLINET (Oltingue) un curieux jeu où une anche battante était placée dans un tuyau en bois à section rectangulaire surmonté d'un autre tuyau en étain de forme tronconique inverse, fermé en haut, mais comportant une ouverture en forme de secteur. Les " résonateurs " couplés dans lesquels vibraient l'anche représentent précisément une simulation remarquable de l'appareil phonatoire,

..../



qui implique des observations intelligentes de celui-ci.

Toutes ces données, toutes ces connaissances restaient cependant disparates du temps de Mersenne. Elles vont se cristalliser au 18^e siècle, certainement sous l'influence de la mode des automates. Les automates jouant d'un instrument de musique, commandés par un " programme " sur tambour à picots, étaient alors très appréciés. Ah! si on réussissait à faire parler un automate.... Ce désir a dû stimuler de nombreux chercheurs à l'époque, et la simulation mécanique de la parole sera le pôle vers lequel vont converger de nombreuses recherches.

Dans une publication récente, Mme IETTAS fait l'historique du problème (bib.1); nous en extrayons quelques noms et dates significatifs :

1780 KRATZENSTEIN, en Pologne, publie un mémoire sur la " Naissance et la formation des voyelles. "

1783 l'abbé MICAL présente à l'Académie des Sciences deux " têtes parlantes " dont l'une disait : " le roi a donné la paix à l'Europe ! " et l'autre répondait : " la paix fait le bonheur des peuples ". Ces têtes comportaient une soufflerie, une série de cavités variables, des soupapes, etc... le tout actionné par un programme sur cylindre. Le rapport signé de noms illustres (LE ROY, de MILLY, LAPLACE, PERIER, VIC D'AZIR et LAVOISIER) précise que la machine prononce bien les voyelles mais que certaines consonnes ont besoin d'être améliorées car les phrases " ne sont pas distinctement prononcées dans toutes leurs parties. "....

1791 le Baron von KEMPELEN populaire pour son automate, le " joueur d'échecs ", qui défraya la chronique de l'époque, publie un important ouvrage à Vienne (en français et en allemand); c'est :

" Le mécanisme de la parole, suivi d'une description d'une machine parlante par M. de KEMPELEN; Conseiller aulique actuel de sa majesté l'Empereur Roi ".

Cet ouvrage résume un travail acharné de recherche qui durerait depuis plus de 20 ans... Le " parleur " que von KEMPELEN avait réalisé suscita un vif intérêt; l'un des modèles réalisés existe encore au musée de MUNICH.

Von KEMPELEN réussit certainement à simuler correctement de nombreux mots et de petites phrases; son ouvrage comporte de nombreux dessins très précis et toutes les explications nécessaires pour réaliser et faire fonctionner la machine.... Il fallait certainement apprendre à " jouer " de cet " instrument " et les opérations étaient probablement assez délicates. Mais les dispositions adoptées par Von KEMPELEN montrent que celui-ci avait très bien compris l'importance et le rôle des transitions phonétiques

.../

pour simuler la parole articulée. Cependant comme tout le monde parle de cette machine sans être très bien fixé sur son rendement réel, l'un d'entre nous, J.S. LIENARD, a décidé de reconstruire le parleur de von KEMPELEN. Ceci permettra de savoir ce qu'il en est et de faire, avec les moyens d'analyse dont nous disposons au laboratoire, de nombreuses observations intéressantes auxquelles une réunion ultérieure du GAM sera consacrée.

Après Von KEMPELEN, de nombreux chercheurs vont se passionner pour cette question; parmi eux des savants tels que WHEATSTONE et d'autres se contenteront de copier le modèle de von KEMPELEN. Les résultats ne pourront guère dépasser ceux qu'il avait obtenus von KEMPELEN parce que les moyens mécaniques utilisés restent sommaires. HELMHOLTZ réalisera bien la synthèse de certaines voyelles à l'aide de diapasons mais cela n'apporte pratiquement rien de nouveau au problème et il faudra attendre les possibilités offertes par l'électro-acoustique pour avancer la question.

Dès 1922 John Q. STEWART réalise une simulation électronique que FLETCHER perfectionnera (1924). Des idées nouvelles se font jour à la suite de l'invention du VODER et du VOCODER, juste avant la guerre (1939); ce dernier appareil permettait de faire simultanément l'analyse et la synthèse d'une parole particulière, mais intelligible. Dès lors, et surtout après guerre, on vit apparaître une floraison de publications sur la parole dont l'étude était désormais facilitée par l'apparition d'analyseurs bien adaptés à ce problème (type sonagraphe). De nombreuses théories nouvelles virent le jour : théorie des voyelles, des formants, du locus, du fondamental (pitch) etc... La théorie de l'information apporta de son côté des éléments nouveaux dans le domaine de la communication des messages et de nombreux laboratoires bien outillés développent actuellement ces problèmes.

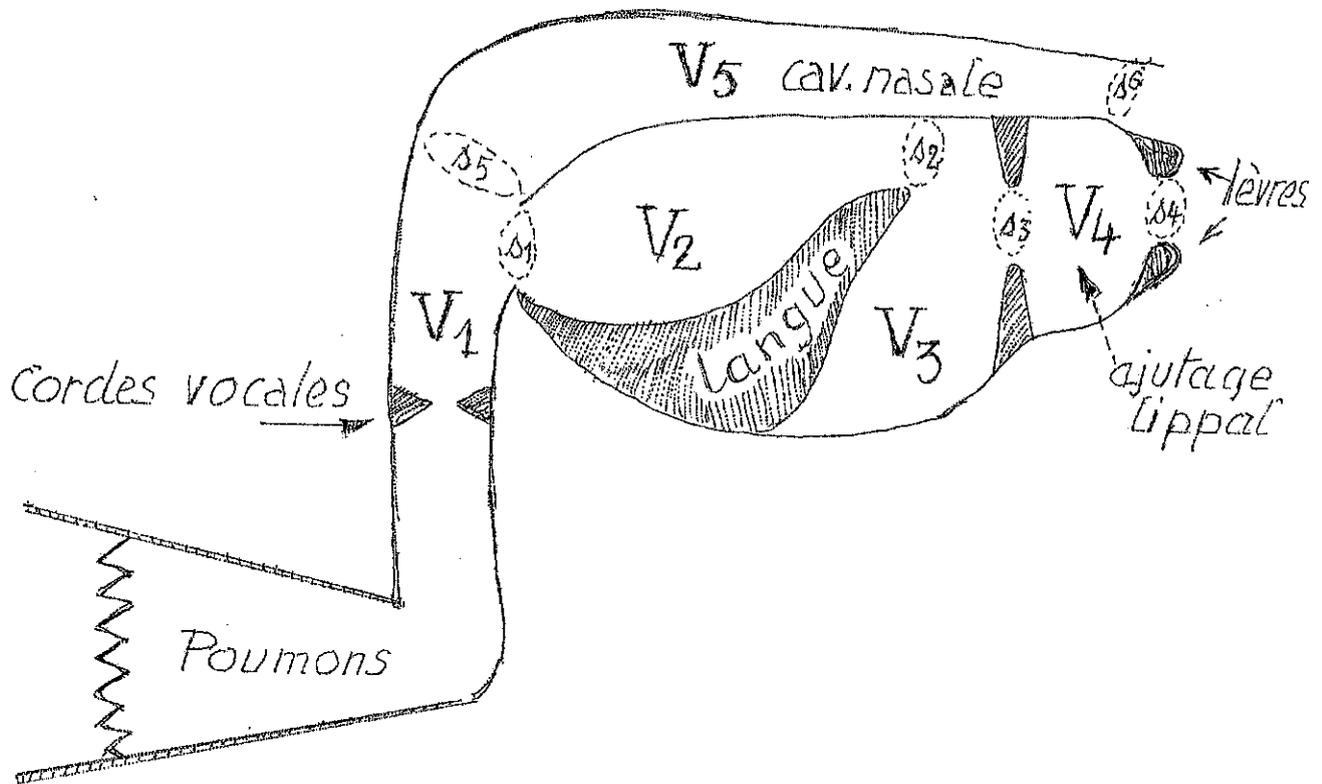
Diverses circonstances nous ont conduit à nous intéresser à ces questions qui rentrent très naturellement dans le cadre général de nos recherches. En effet, musique, bruit, parole ne sont que des aspects différents d'un seul et même problème : étude du fonctionnement de générateurs acoustiques, de la structure physique des signaux rayonnés, des canaux de communication en tant que destructeurs des signaux, de la perception et de l'intégration des signaux par l'homme. En bref, la parole ne représente qu'un cas particulier de communication de messages, se distinguant des autres uniquement par des particularités de structure et de fonctionnement de l'"instrument" de musique qui sert à la fabriquer et par la nature particulière du message du point de vue informatif.

II. SYSTEME PHONATOIRE ET STRUCTURE PHYSIQUE DE LA PAROLE

1°) LE SYSTEME PHONATOIRE ET SON FONCTIONNEMENT.

L'anatomie du système phonatoire est bien connue et notre intention est seulement de donner un schéma de fonctionnement permettant de préciser plus clairement un certain nombre de points importants.

Ce système comporte (fig. 1)



- une soufflerie (les poumons) qui fournit l'énergie nécessaire sous forme d'air comprimé dont la modulation conditionne les ondes acoustiques rayonnées par le système.

...../

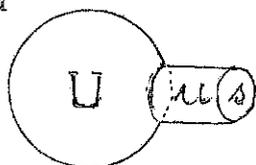
- une anche double, musculaire qui fournit, comme tout dispositif similaire, un spectre de raies très riche comportant de nombreux harmoniques (30 à 40 couramment.) Les anciennes théories admettaient que ces anches vibraient sous l'effet de l'air comprimé, comme une anche de basson; récemment on a montré que la commande de fréquence de vibration provenait du cerveau. En fait, il est évident que dans ce dernier cas l'écoulement aérien réagit sur le fonctionnement du dispositif. L'important est de retenir le rôle de générateur de sons harmoniques dans le cas présent.

- une série de cavités à parois membraneuses, plus ou moins lisses, toujours humides. Les formes et les dimensions de ces cavités ainsi que les sections de leurs ouvertures peuvent être modifiées à volonté, de façon autonome ou simultanée, entre des positions-limites imposées par l'anatomie et la physiologie des organes phonatoires. On observe que les cavités V2 et V3 sont complémentaires, et, à la limite (langue plaquée vers le bas), la cavité V3 disparaît.

Ces cavités jouent le rôle de résonateurs couplés en série - sauf la cavité nasale V5 qui est en parallèle.

En raison de l'amortissement, seuls les partiels 1 (fondamentaux) jouent un rôle notable dans la voix normale. On sait qu'en première approximation, la fréquence d'un résonateur de ce type (fig.3) est égale à

$$N = \frac{V \cdot s}{2\pi \sqrt{U \cdot u}}$$



N = fréquence

V = vitesse du son

s = section de l'ouverture

U = volume du résonateur

u = volume de l'ajutage.

Ces fréquences de résonance déterminent les "formants" de la parole. Du fait de l'amortissement du système, ceux-ci se présentent comme des bandes larges dont la situation dans le diagramme fréquence-temps (sonagramme) est fonction des variables définies plus haut. Le locuteur dispose, pour chacun d'eux, d'un "champ de liberté" qu'il exploite au mieux.

La formule ci-dessus est très théorique; pour des résonateurs compliqués et couplés, seules les machines à calculer modernes pourraient venir à bout du problème. Mais elle a le mérite de mettre en évidence les variables en cause; elle permet aussi de se faire une idée du fonctionnement du système. Ainsi on notera en particulier que la section de l'ouverture et les volumes agissent en sens inverse. On peut donc obtenir une certaine "hauteur" de formant pour de nombreuses combinaisons

...../

entre s et U; mais il est évident que chaque ouverture commande deux résonateurs dans le système considéré : modifier un élément de ce système revient à modifier l'ensemble qui doit être considéré comme un " tout " où chaque partie réagit sur l'ensemble.

On retiendra encore que les différences entre individus peuvent être considérables (enfants, femmes, hommes), mais que les systèmes phonatoires humains sont statistiquement semblables.

2°) LA STRUCTURE PHYSIQUE DES SIGNAUX PHONÉTIQUES.



Le système phonatoire précédent est normalement excité par deux phénomènes différents, à savoir le spectre de raies délivré par les cordes vocales et le bruit d'écoulement gazeux. Ce dernier comporte plusieurs aspects : bruit normal de frottement d'un gaz contre les parois (intense, en particulier, au voisinage des ouvertures lorsqu'elles sont petites) et bruits de détentes gazeuses (explosions vers l'extérieur ou implosions vers l'intérieur d'une cavité). Explosions et implosions sont généralement très brèves, de l'ordre de quelques millisecondes.

Les signaux acoustiques rayonnés par l'ensemble sont alors de deux types distincts :

- Des phénomènes quasi-stationnaires, lorsque le système phonatoire est bloqué dans une position fixe. On peut ainsi produire les voyelles et un certain nombre de consonnes (sifflantes, chuintantes etc...). Les phénomènes réalisés ainsi correspondent à un ordre de position des divers organes en présence : langue, dents, lèvres etc...

- Des phénomènes évolutifs. On observe facilement que, lors de la parole normale, le système phonatoire n'est jamais immobile. La parole n'est donc pas une succession de phénomènes stationnaires, mais un enchaînement de mouvements, donc un phénomène transitoire qui traduit un ordre de mouvements, réglé par un programme. Acoustiquement il est donc impossible de décrire la parole avec les spectrographes utilisant la méthode de la boucle fermée; le seul type d'analyseur qui conviendrait actuellement le sonagraphe. Cet appareil et les documents qu'il fournit sont maintenant bien connus de tous, et nous avons maintes fois attiré l'attention sur son adéquation pour étudier les signaux évolutifs. L'intérêt qu'il présente réside dans le fait qu'il fournit sur un document unique l'évolution simultanée de l'intensité et de la fréquence en fonction du temps et que ce document peut être " appréhendé " visuellement en bloc, comme un tout, ce qui permet, moyennant un apprentis-

sage, de faire aisément la corrélation entre audition et document. Rappelons que la précision de l'appareillage actuel est largement suffisante puisqu'on retrouve sur le document tout ce que "l'oreille" peut saisir.

En bref, le sonographe nous donne une description physique suffisante, et cela, d'autant plus que la parole est un phénomène flou par définition en raison de l'amortissement du système phonatoire. Le sonogramme peut facilement être "quantifié" pour l'étude de détail, c'est-à-dire découpé en petites surfaces élémentaires de perception que nous appelons "quantas acoustiques". Chacun de ces quantas représente une certaine quantité d'énergie rayonnée par le système phonatoire, mais ils ne jouent pas tous le même rôle. Certains sont inutiles pour l'intelligibilité du discours normal, d'autres sont indispensables; on peut s'en assurer facilement par des expériences de filtrage classiques où l'on détruit systématiquement telle ou telle partie du phénomène acoustique. Citons entre autres, une expérience qui nous a semblé significative.

En octobre 1964, M. GEISSERT présentait à l'Institut des Sourds Muets une thèse très intéressante. Il s'agissait d'un texte littéraire (la chèvre de Monsieur Séguin de A. DAUDET) préalablement enregistré sur bande et que l'on avait découpé en fragments de 2,4 secondes. Chaque fragment avait été analysé au sonographe et les sonogrammes furent collés bout à bout. Le rouleau de sonogrammes fut ensuite filmé. Ainsi on pouvait suivre simultanément le texte récité et l'analyse sonographiqu.

Nous avons au laboratoire une longue expérience dans le dépouillement des sonogrammes mais il était impossible pour nous de faire la corrélation entre image et parole; c'est tout juste si au passage nous reconnaissons de temps à autre un signal très caractéristique comme les sifflantes ou les chuintantes: l'image défilait trop rapidement, en d'autres termes elle véhiculait beaucoup trop d'information simultanée pour que nous soyons capable de l'intégrer. Le langage parlé contient une grande redondance et nous étions incapables d'extraire l'essentiel du document visuel.

Ceci nous a conduit à nous intéresser de près à la nature informative du message parlé, et à rechercher des moyens techniques pour opérer un clivage entre les divers types d'information contenu dans la parole.

III. NATURE INFORMATIVE DU MESSAGE PARLE

La voix, ensemble complexe de signaux acoustiques possibles avec les éléments du système phonatoire humain reliés entre eux, ne devient parole que si elle véhicule une information intelligible pour un récepteur humain. Cela suppose au préalable un " code " commun à l'émetteur et au récepteur. L'association entre un " signe " acoustique et une signification est purement arbitraire, et l'existence d'innombrables langues humaines suffit à le montrer. Que l'on puisse retrouver dans les diverses langues quelques similitudes n'a rien d'étonnant : parmi les signaux acoustiques que l'on peut réaliser avec l'appareil phonatoire tel qu'il est, certains sont faciles à produire et s'imposent en vertu du principe du moindre effort (dans ce sens, le mot " maman " est " naturel "... d'autres le sont moins. On pourrait faire sur ce point une étude très intéressante; le " code " lui même est purement conventionnel et nécessite un apprentissage : il n'y a aucun rapport évident entre structure physique de la parole et signification du message.

Quoiqu'il en soit, pour communiquer le message à distance il faut passer par un canal. Il n'est pas douteux que les langues actuelles se sont élaborées graduellement en tenant implicitement compte des caractéristiques du canal " naturel ", c'est-à-dire du canal aérien direct. Tout canal introduit des distorsions et du bruit de fond; si on désire que la signification parvienne au récepteur en restant intelligible il faut donc que le message résiste à la destruction due au canal. Dans la parole normale, cette destruction provient du fait que l'air absorbe graduellement les vibrations aériennes, en commençant par l'aigu. D'autre part, les bruits de fond très variés de la nature détruisent tantôt telle région fréquentielle, tantôt telle autre; nous avons montré ailleurs le mécanisme de l'émergence d'une forme sur le fond.... (bib.2). Dans ces conditions, une redondance considérable est nécessaire, mais il est évident que cette redondance n'est utile que pour la " portée " à distance, et dans le cas de l'existence d'un bruit de fond. Il existe donc dans le signal physique de la parole une " information à l'état pur ".

En fait, l'expérience montre que la parole véhicule trois types différents d'information différents; ce sont :

- l'information parasitaire. Elle comporte des bruits de fonctionnement sans signification provenant de l'appareil phonatoire. Ce sont, par exemple, des bruits fixes totalement prévisibles, tels des sifflements, et ne véhiculant par suite que de l'information annexe sans rapport avec le message véritable (cet homme a la bronchite ... etc), ou encore des bruits erratiques totalement imprévisibles (bruits de salive, etc...) sans intérêt du point de vue qui nous occupe. En bref, ce sont toutes les composantes acoustiques qui ne sont pas et ne peuvent pas être réglés systématiquement et volontairement par le locuteur.

...../

- l'information esthétique. Nous portons sur la voix d'un locuteur donné des jugements de qualité esthétique : elle est " chaude ", " timbrée ", " sèche ", " aigre ", etc... Ces qualités sont liées à la structure particulière, à l'amortissement de son système phonatoire, au tonus de son système nerveux; elles représentent ce que le locuteur a de différent d'un autre et nous permettent de le distinguer parmi d'autres. Cette information présente un certain intérêt dans des cas bien définis; dans le chant, par exemple elle joue un rôle primordial; mais le véritable problème de la parole n'est manifestement pas là.

- l'information sémantique. Elle représente dans le discours ce que l'on peut traduire en allemand, en anglais.... L'information sémantique c'est la parole. Pour étudier l'information sémantique il faut trouver un moyen de l'isoler. C'est dans ce sens que nous avons alors développé nos recherches.

IV. L'EXTRACTION DE L'INFORMATION SEMANTIQUE ET SON INTEGRATION PAR L'HOMME

Nous ne considérerons ici que les langues européennes et plus spécialement la langue française, à l'exclusion des langues à tons où le problème se pose différemment. Nous avons une étude en cours sur ce dernier sujet, qui fera l'objet d'une publication ultérieure.

Il est bien connu de tous que l'on peut éliminer une partie de la redondance à l'aide de méthodes classiques de filtrage; l'expérience montre qu'on peut supprimer tout ce qui dépasse 3500 Hz sans nuire pratiquement à la reconnaissance des mots. Mais diverses expériences nous ont montré que la bande passante utile véhicule encore une forte proportion d'information parasitaire et esthétique, qu'il est impossible d'éliminer par filtrage, car on détruit alors également l'information sémantique.

Deux observations nous ont conduit vers une solution simple et efficace :

1°) L'EXPERIENCE DE LA GUIMBARDE.

Elle a fait naguère l'objet d'une communication au GALP (bib.3). Nous avons attiré alors l'attention sur l'intérêt de ce modeste instrument de musique populaire, constitué d'une simple lame d'acier fixées sur un cadre en fer, et que l'on pince avec un doigt en tenant le cadre contre les dents entrouvertes.

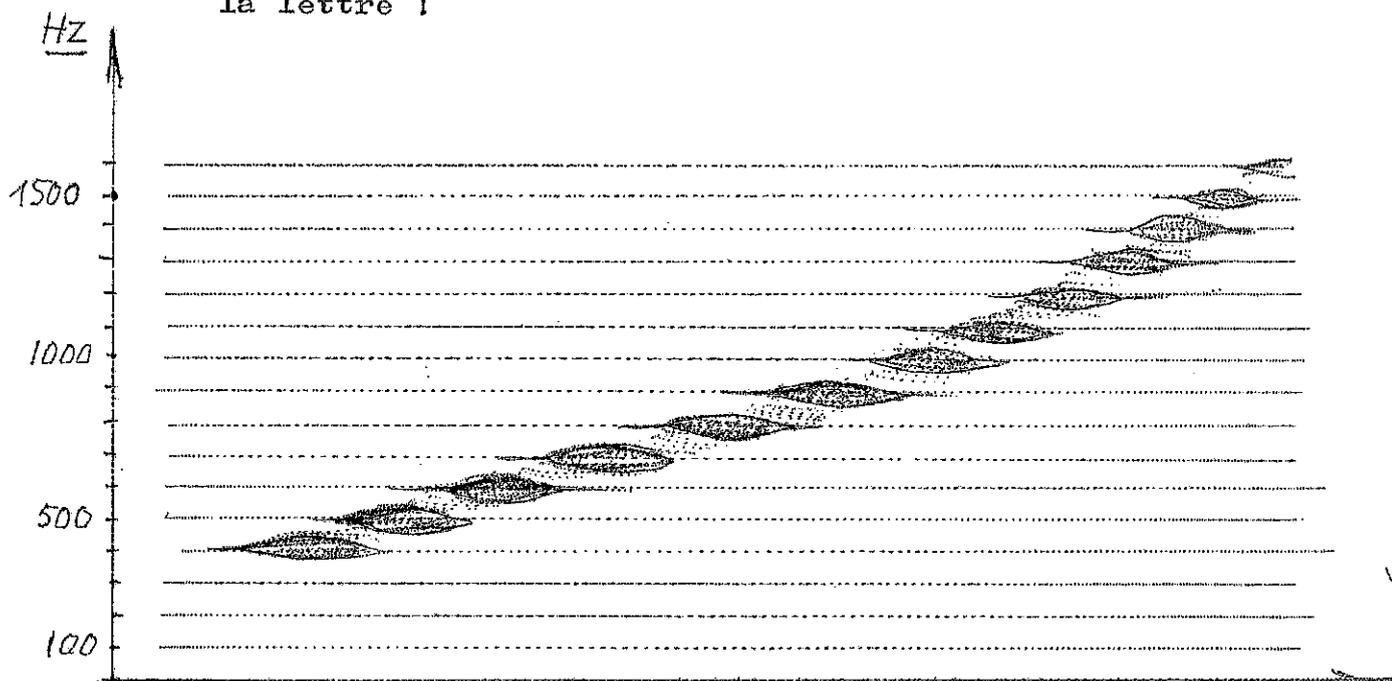
...../

Si on analyse le signal de cette lame, on observe qu'elle rayonne un spectre extrêmement riche de partiels quasi-harmoniques (30 à 40 partiels au moins....), dont le fondamental varie selon les modèles entre 50 et 100 Hz environ.

Le niveau du signal délivré est très faible : on entend un son très grave correspondant à peu près au fondamental de la lame.

Mais si on associe cette lame à un résonateur convenablement accordé, on entend le son au loin : il y a " amplification par résonance " .

Or dans le jeu normal de l'instrument, le musicien utilise comme résonateur sa cavité buccale, qu'il peut accorder à volonté entre 500 et 2000 Hz environ. Si on pince l'instrument posé sur les dents et si on diminue graduellement la cavité buccale, on entend nettement deux phénomènes associés, distincts : un son grave fixe (pédale) provenant du fondamental invariable de la lame et " quelque chose qui monte " Ce quelque chose est tout simplement le " formant " de la cavité buccale; le phénomène est extrêmement net sur le sonagramme : chaque fois que la cavité buccale " passe " sur l'un des partiels, celui-ci est amplifié par résonance (fig.2). Les joueurs de guimbarde réalisaient donc une véritable "musique de formants " bien avant la lettre !



*Le fondamental de la lame est 100 Hz ici.
Les raies horizontales sont les partiels
Le résonateur buccal amplifie successivement les partiels.*

Figure 2

...../

Cette expérience nous a donné l'idée d'en faire une autre, liée aux préoccupations que nous avons alors au sujet de la parole. On procède de la façon suivante :

On déconnecte la soufflerie de l'appareil phonatoire ainsi que les cordes vocales, pour ne conserver que les mouvements du système phonatoire - donc les modifications de formes et de dimensions du système de résonateurs couplés; on parle en " muet "

On place alors la guimbarde entre les dents, et.... on a la surprise d'entendre une parole semi-synthétique parfaitement intelligible, qui ressemble à s'y méprendre à celle d'un vocoder.... Il n'y a là finalement rien d'étonnant ; un sonagramme de vocoder et un sonagramme de parole-guimbarde sont strictement identiques

Cette expérience montre deux choses importantes :

a) l'information sémantique est surtout déterminée par l'ordre de mouvement des cavités antérieures du système phonatoire.

b) le spectre de raies ne véhicule aucune information sémantique notable puisqu'il reste strictement fixe.

Les conséquences de ces observations sont d'importance capitale comme on verra par la suite. Mais étant donné l'importance que l'on attache généralement au spectre de raies en parole, et plus spécialement au fondamental - sur lequel de nombreux laboratoires font actuellement des recherches systématiques - il est nécessaire de préciser le rôle du spectre de raies.

2°) LE RÔLE DU SPECTRE DE RAIES.

Il est évident que la parole normale, élaborée depuis des siècles, s'est épurée de tout ce qui n'était pas strictement utile à quelque chose. Il faut donc trouver une justification à l'existence du spectre de raies. Nous proposons l'explication suivante : le spectre de raies joue uniquement le rôle " d'onde porteuse " permettant, pour une énergie disponible donnée, de porter un message acoustique le plus loins possible. Pour le montrer, considérons un bruit blanc de 80 dB composé sur le sonagramme de quantas acoustiques répartis au hasard pendant un temps donnée (fig.3a); ce bruit "porte" à une certaine distance dans des conditions données. Mais regroupons maintenant les " quantas acoustiques ", les " grains d'énergie "

..../

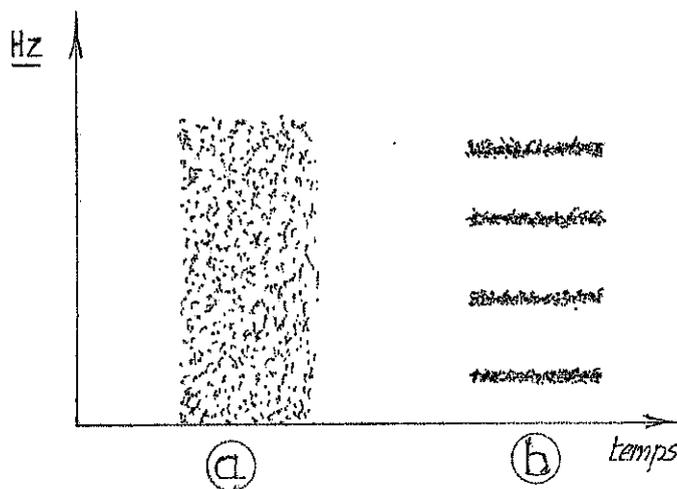


Figure 3

autour de lignes horizontales (b); nous avons constitué ainsi un " spectre de raies " et il est bien évident que l'énergie concentrée sur une raie est très supérieure à celle de la région correspondante du bruit blanc... Le spectre de raies portera donc plus loin...

Il faut donc bien préciser que dans la parole normale les cordes vocales jouent simultanément un double rôle :

- elles font monter et descendre la voix, la rendant ainsi moins monotone, moins fatigante, moins ennuyeuse etc... mais c'est là uniquement de l'information esthétique.
- elles conditionnent la portée de la voix.

Que le spectre de raies véhicule obligatoirement de l'information sémantique est évident; mais c'est une conséquence indirecte provenant du fait qu'il est modulé par les mouvements des cavités. Mais qu'il soit inutile à l'information sémantique est facile à montrer, et c'est ce qui nous a conduit aux expériences sur la voix chuchotée.

3°) LA VOIX CHUCHOTÉE.

Elle a intéressé de nombreux chercheurs, dès l'apparition du visible speech (prototype du sonographe), et plusieurs publications y font allusion, dont certaines très intéressantes (bib.4). Il nous est cependant apparu qu'on n'avait pas tiré toutes les conséquences des observations faites à ce sujet.

Nous avons tout naturellement été conduits à nous intéresser à la parole chuchotée à la suite de l'expérience de la guimbarde : celle-ci nous avait montré que le spectre de raies ne véhiculait pas l'information sémantique, mais il était difficile de prononcer correctement tous les mots en raison de la gêne évidente que provoque le cadre de la guimbarde pour exécuter normalement les mouvements habituels de l'appareil phonatoire. C'est alors que nous avons pensé à la voix chuchotée.

Celle-ci, par définition, ne contient pas de raies puisque les cordes vocales sont déconnectées. On sait aussi que de ce fait elle ne porte pas loin, avantage notable à l'école ou au lycée... Par contre on vérifie sans difficulté qu'elle est tota-

...../

lement intelligible. La conclusion est donc immédiate : la voix chuchotée traduit intégralement l'ordre de mouvement de l'appareil phonatoire dans les conditions normales de prononciation; elle véhicule donc intégralement l'information sémantique. Son étude systématique nous a livré la clef de ce que nous cherchions : un moyen d'extraire l'information sémantique à l'état le plus pur possible.

Nous avons rapidement observé que même la parole chuchotée comportait encore un fort taux de redondance; on peut pratiquement couper tout ce qui est au-dessus de 3000 Hz et on vérifie que l'intelligibilité reste totale. Nous avons ensuite cherché à augmenter le contraste des images sonographiques par des " cuisines " acoustiques appropriées et finalement il ne restait sur les sonagrammes que des traces. Mais, et l'observation est capitale, ces traces présentent tous les caractères de formes, au sens de la Gestalttheorie.

Dès lors l'idée s'est faite que comprendre la parole c'est percevoir, reconnaître et intégrer des formes.

La reconnaissance des formes pose des problèmes nombreux et difficiles que plusieurs laboratoires spécialisés étudient actuellement à l'aide de moyens puissants, ordinateurs entre autres. Ces questions nous concernent directement, et nous travaillons depuis longtemps sur un modèle de fonctionnement de la perception qui fera l'objet d'une publication ultérieure. Nous proposons un schéma très simplifié qui nous permettra cependant de poser le problème clairement : (fig.4)

Un locuteur prononce un mot ou une phrase; le système phonatoire transforme l'idée en vibrations acoustiques de l'air ambiant; celles-ci sont captées par l'oreille qui transforme les vibrations aériennes en vibrations électriques qui peuvent alors être traitées sous cette forme par le " laboratoire " électro-chimique que représente notre cerveau.

Il se produit d'abord une projection, une visualisation du signal (que nous admettrons comme similaire à un sonagramme pour simplifier les choses.) Dès notre naissance nous voyons ainsi défiler des images que nous stockons dans une " mémoire des signes ". Pour ne pas encombrer celle-ci, nous procédons de temps à autre à une opération mentale de réflexion consistant à comparer les formes similaires, à en extraire les invariants et à ne conserver finalement que les images qui seront d'autant plus épurées, stéréotypées, que le même signal reviendra plus souvent sur l'écran de projection.

Dans une deuxième mémoire, la " mémoire des significations " nous stockons de même, par conditionnement, des " cartes " portant ^{chacune} un signe et sa signification (arbitraire, mais identique dans une même société).

..../

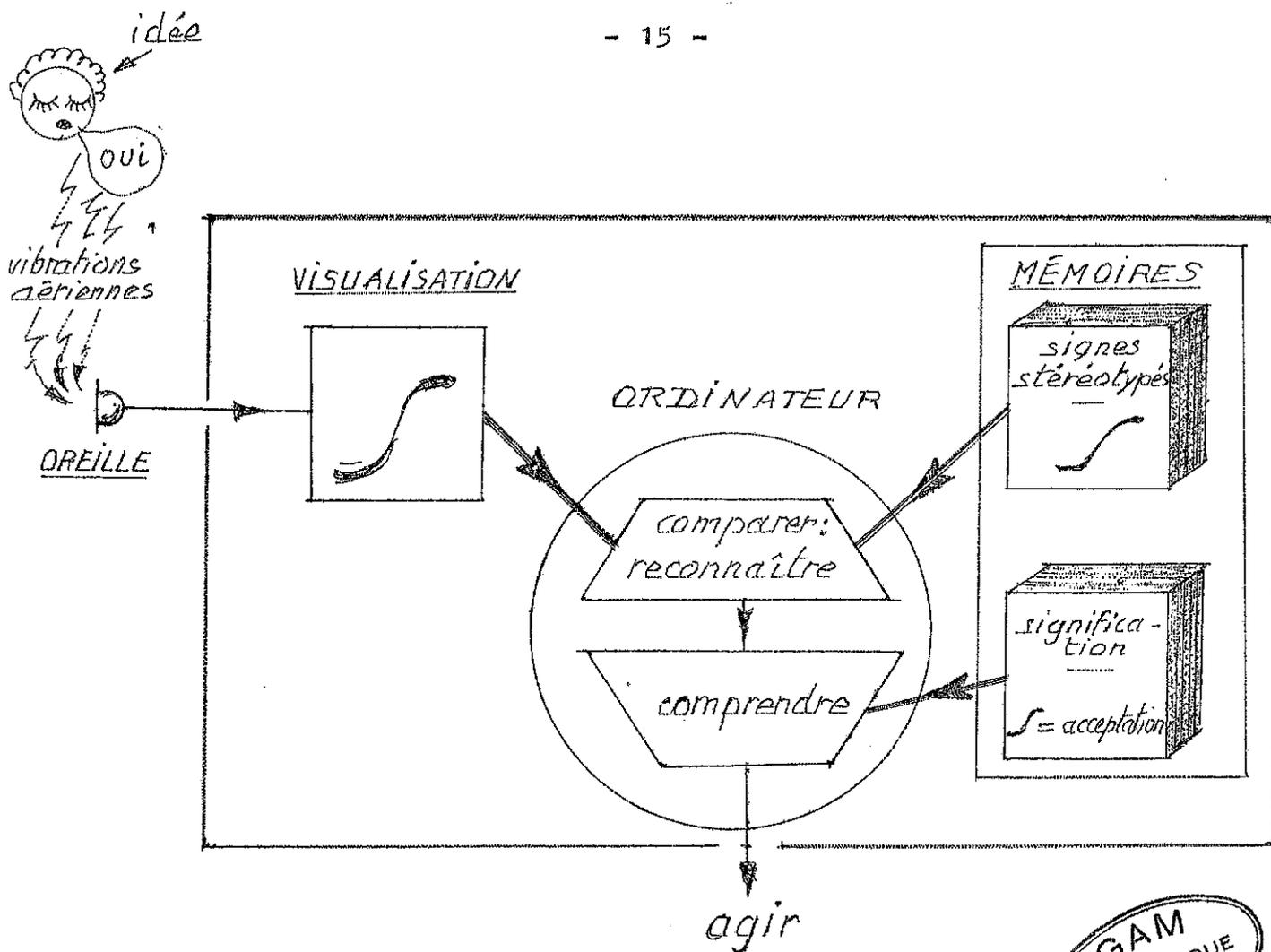


Figure 4

Comprendre la parole comporte dès lors deux opérations distinctes :

- reconnaître le signe en puisant dans la mémoire des signes
- comprendre le signe en puisant dans la mémoire des significations.

Or nous venons de voir que l'information sémantique de la parole est représentée par les " formes " que l'on extrait de la parole chuchotée. Nous sommes donc tout naturellement conduits à nous intéresser aux travaux des psychologues sur la perception des formes et tout spécialement à la **GESTALTTHÉORIE** qui s'est révélée, depuis 50 ans comme une des théories les plus fructueuses, dont les lois s'appliquent à peu près intégralement dans le cas qui nous intéresse ici. Les éléments de cette théorie sont maintenant bien connus et de très nombreuses publications dans toutes les langues ont été faites sur ce

..../

sujet depuis les travaux de von EHRENFELS à la fin du siècle dernier. Il n'est cependant pas inutile de reprendre les principaux points qui nous concernent ici.

V. GESTALTTHÉORIE ET PAROLE

Les gestaltistes considèrent les formes comme des unités, perçues en bloc; Les messages du monde extérieur nous parviennent sous l'aspect de formes, qu'elles soient auditives, visuelles, tactiles ou autres. En fait notre schéma précédent (fig.4) est tout à fait général de ce point de vue. Une sensation comporte trois dimensions : l'intensité, la durée et la " composition structurale " spectrale par exemple. Il n'est donc pas étonnant que la gestalt-théorie s'applique à toutes nos perceptions.

Parmi les lois que la gestalttheorie a élaborées nous retiendrons les suivantes :

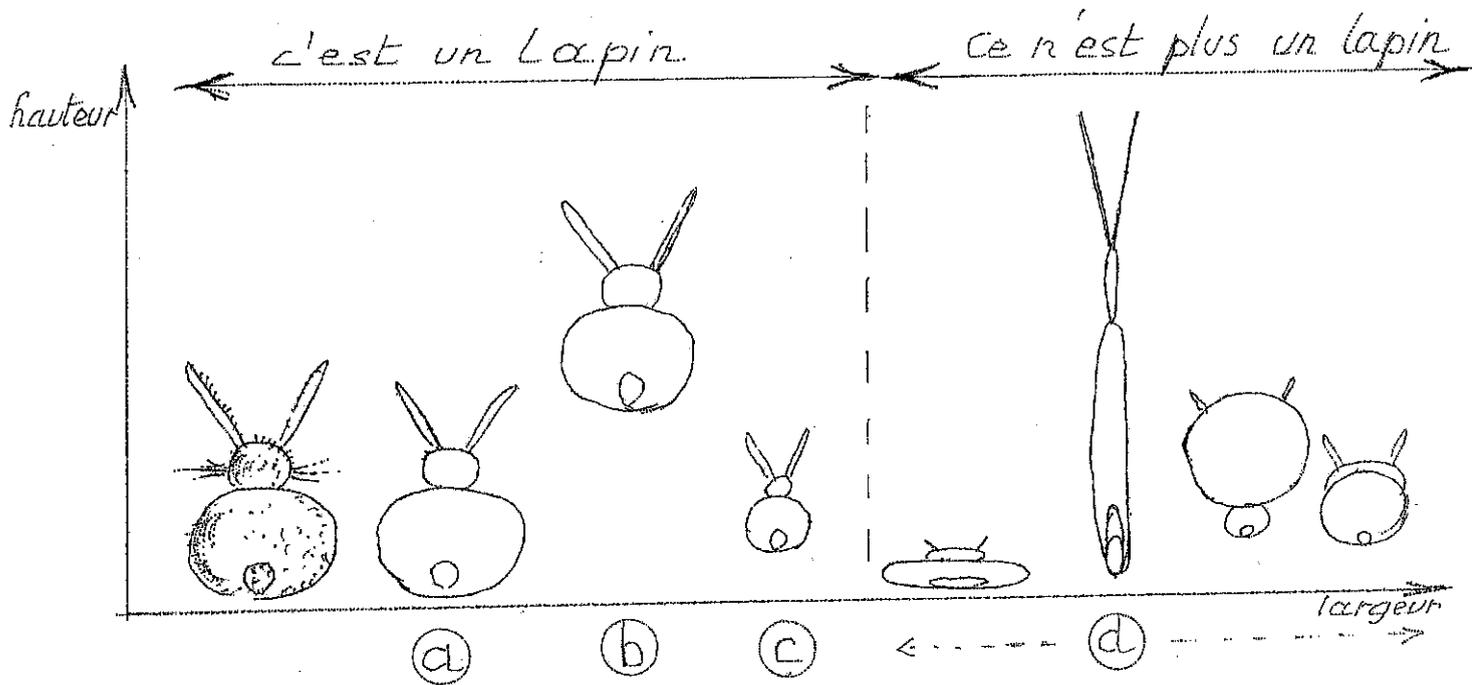


Figure 5

- 1°) Il existe des formes fortes et des formes faibles. Plus une forme est forte, plus elle est difficile à détruire et plus elle est facile à styliser, à stéréotyper, à suggérer à l'aide de peu d'éléments significatifs. Prenons un cas précis de forme visuelle : le lapin (fig.5). Depuis notre enfance nous avons vu cette forme tant de fois que nous en possédons dans la mé-

...../

moire une forme stéréotypée extrêmement épurée ; nous reconnaissons l'animal sans ambiguïté sous l'aspect du croquis simplifié (a). En fait nous projetons le stéréotype sur le signal actuel : si la corrélation est totale nous " reconnaissons " la forme. Cet exemple va nous permettre de préciser simplement d'autres lois de la gestalttheorie :

- 2°) Une forme est transposable sans cesser d'être identifiable. Ainsi nous pouvons placer notre dessin plus haut ou plus bas dans les coordonnées choisies (b); nous pouvons aussi en réduire les dimensions (c) : la forme reste toujours reconnaissable. L'expérience nous montrera qu'il en est strictement ainsi pour les " formes " de la parole chuchotée-filtrée; c'est le cas, en particulier des voix d'enfant
- 3°) Une forme est anamorphosable(d) nous pouvons en modifier les proportions dans une très large mesure et d'autant plus que la forme est plus forte. Mais la reconnaissance mentale est d'autant plus facile :

- que les formes contenues dans notre mémoire seront plus stéréotypées - c'est-à-dire perçues plus de fois,

- que " l'ordinateur " du récepteur sera plus " puissant " (autrement dit que l'individu sera plus intelligent).

- 4°) Une forme n'est pas la somme de ses éléments, mais une fonction de plusieurs variables. Avec cette proposition nous entrons véritablement dans le coeur du sujet. La Gestalttheorie s'est établie comme une réaction contre la théorie analytique de la perception, qui prétendait expliquer les formes en les découpant en éléments que l'on étudiait isolément. Or on procède ainsi en phonétique si on découpe les mots en phonèmes. L'élément, en parole, c'est le mot, perçu comme forme globale. L'expérience montre que le langage parlé n'est pas constitué de phonèmes additionnés les uns aux autres, précisément parce que chaque phonème change de forme selon la situation et la fonction qu'il occupe, en d'autres termes, selon ce qui précède et ce qui suit. Divers auteurs ont déjà attiré l'attention sur ce problème (Théorie du locus de DELATTRE; travaux sur les transitions phonétiques de LAFON, etc...) mais on ne semble jamais avoir insisté sur le fait qu'un mot est une totalité, une forme globale originale qu'il est impossible de confondre avec une autre et dont les sonagrammes de parole chuchotée filtrée donnent une représentation extrêmement significative de ce point de vue.

On vérifie facilement par ce procédé que le mot " oui " n'est pas du tout la somme de " ou " et de " i " mais une forme en S plus ou moins couchée (fig.6); de même " pa " n'est pas la somme de " p " et de " a ".

..../

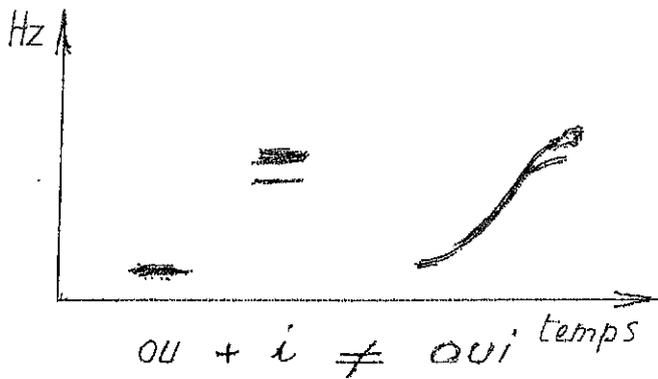


figure 6

Dès lors, la conclusion est immédiate : le mot est la " forme " fondamentale de la parole. Cette forme est la traduction de l'ordre de mouvement de l'appareil phonatoire; on peut la découper, non en phonèmes mais en formes élémentaires que nous appellerons ELEMENTS PHONÉTIQUES. Un élément phonétique est l'image du mouvement entre deux phonèmes Dans un mot les éléments phonétiques se raccordent automatiquement comme se raccordent par la force des choses, les mouvements de l'appareil phonatoire, en allant d'une position limite vers une autre.

En d'autres termes, le mot " parole " n'est pas l'association des 6 phonèmes " p, a, r, o, l, e " mais la fonction des 5 variables (les éléments phonétiques) pa, ar, ro, ol, le. Ceux-ci se soudent automatiquement de façon correcte s'ils sont réalisés par une même personne. Mais on peut aussi imaginer des formes stéréotypées normalisées, constituant alors une véritable sténographie sémantique de la parole, sur l'intérêt de laquelle nous reviendrons plus loin.

Les conclusions sont immédiates. En admettant l'existence des éléments phonétiques, il était intéressant d'en faire une étude systématique, que nous avons entreprise dès lors, et qui nous a conduit à établir un dictionnaire des éléments phonétiques, complet, de la langue française.

VI. ELEMENTS PHONÉTIQUES ET SYNTHÈSE DE LA PAROLE

1°) LE DICTIONNAIRE DES ELEMENTS PHONÉTIQUES.

Nous sommes partis de la sténographie phonémique DUPLOYE que nous connaissons bien. On y trouve 30 signes qui représentent pratiquement tous les phonèmes utilisés en français pour réaliser un langage intelligible. Ces signes sont " cousus " bout à bout pour réaliser des mots, formes visuelles identifiables - qui suivent d'ailleurs rigoureusement tout ce que nous avons dit plus haut au sujet de la perception des formes. Il suffit d'associer les phonèmes qu'ils représentent deux à deux pour obtenir le dictionnaire complet des éléments phonétiques possibles. En bref, on dessine une matrice d'ordre 30 et on porte en ligne et en colonne les 30 phonèmes considérés; on obtient ainsi 900 cases représentant la totalité des éléments phonétiques possibles.

..../.

Il va sans dire que le taux d'occurrence dans un texte donné est variable pour chacun des éléments phonétiques, selon le " style " du locuteur, son vocabulaire, etc.....

Nous n'insisterons pas davantage sur cette question que J.S. LIENARD a longuement développée et dont il nous donne les principaux résultats, en feuillets annexes. Par contre nous allons montrer l'utilité de ce dictionnaire dans le domaine de la synthèse de la parole.

2°) LA SYNTHÈSE DE LA PAROLE.

Pour vérifier nos hypothèses relativement à l'information sémantique, à la parole chuchotée etc..., nous avons commencé par matérialiser les éléments phonétiques définis sous forme de sonagrammes en voix chuchotée-filtrée. Chaque élément phonétique a été enregistré avec nos trois voix (une féminine); les sonagrammes ont été comparés, et nous en avons extrait une " forme normalisée " moyenne. Dès lors il devenait possible de dessiner n'importe quel mot au pinceau et à l'encre, sur papier ordinaire, sans passer par l'analyse préalable du mot au sonographe - puisque les formes stéréotypées ainsi réalisées sont raccordable par définition.

Nous avons donc pu écrire, ainsi, des phrases quelconques. Il s'agissait maintenant de vérifier si ces " dessins " véhiculaient effectivement l'information sémantique. Pour cela il nous fallait un appareil de synthèse.

Cet appareil a été conçu, réalisé, mis au point et utilisé : c'est l'ICOPHONE.

Il s'agit exactement de l'inverse du sonographe dont nous avons d'ailleurs, par motif d'économie, utilisé la partie mécanique et le disque magnétique (fig.7).

On commence par placer la phrase dessinée, autour du tambour tournant. Celui-ci est, comme on sait, solidaire du disque magnétique.

Une cellule photoélectrique entraînée du bas vers le haut explore graduellement le dessin par spires successives de 45 Hz. Chaque fois qu'un point noir passe devant la cellule, un générateur basse fréquence envoie dans la tête d'enregistrement un signal sinusoïdal de fréquence correspondant à celle du " point " qui est lu par la cellule. On ourle, en face d'une génératrice donnée du tambour, l'ensemble des " quantas " acoustiques placés sur cette génératrice (rerecording); quand l'opération est terminée on écoute le signal global reconstitué puis, à titre de contrôle, on tire le sonagramme de la reconstitution.

..../

Nous avons éprouvé de nombreuses difficultés de réglage et de mise au point; cet appareil est très imparfait; mais les échantillons sonores obtenus reconstituaient en parole chuchotée de façon intelligible la phrase dessinée au préalable. Le sonagramme de contrôle montrait que l'appareil de synthèse laissait à désirer; mais les formes des mots restaient reconnaissables visuellement et auditivement....

Fig 7 Schéma de Fonctionnement
de l'ICOPHONE

GAM
BIBLIOTHÈQUE
DE PRÊT

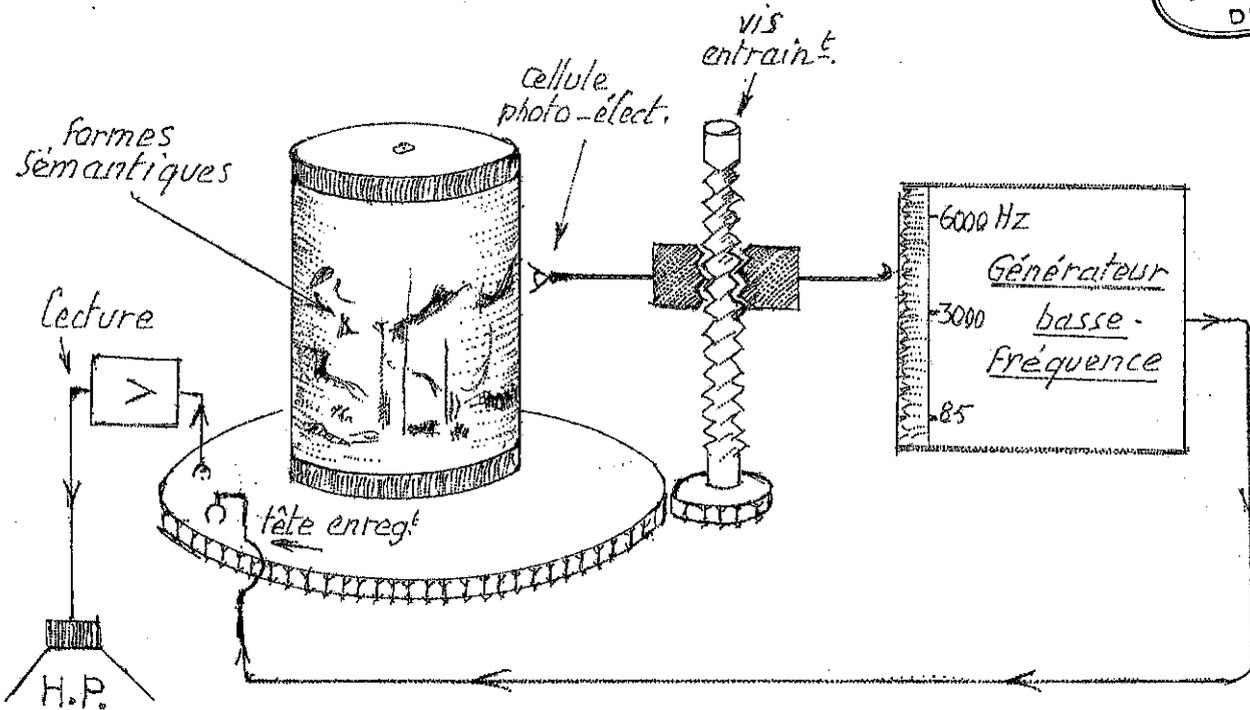
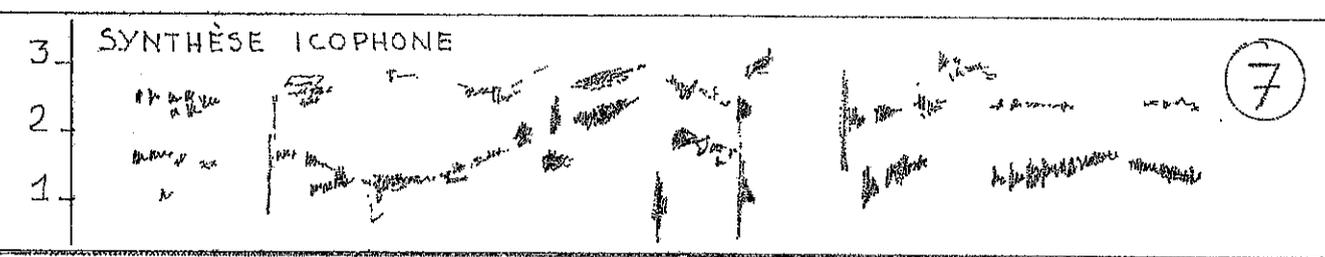
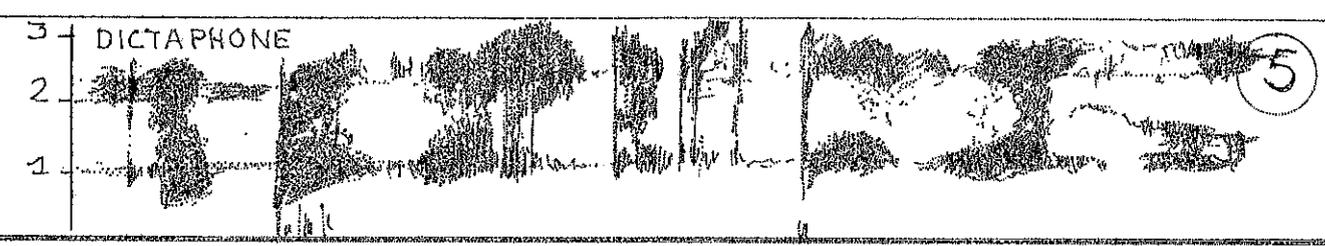
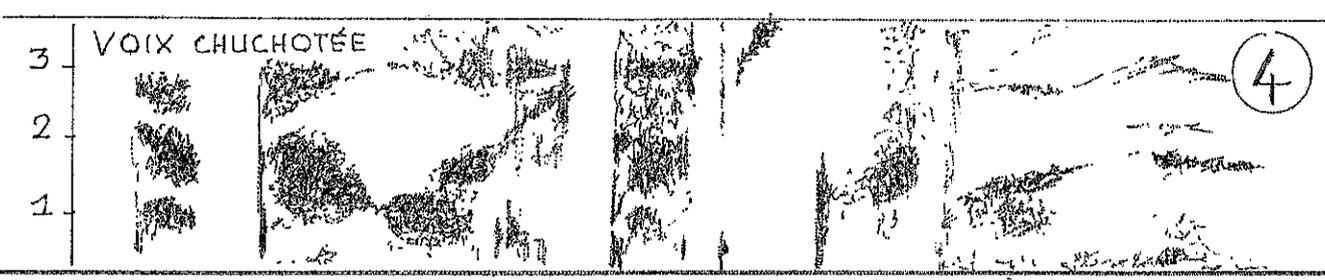
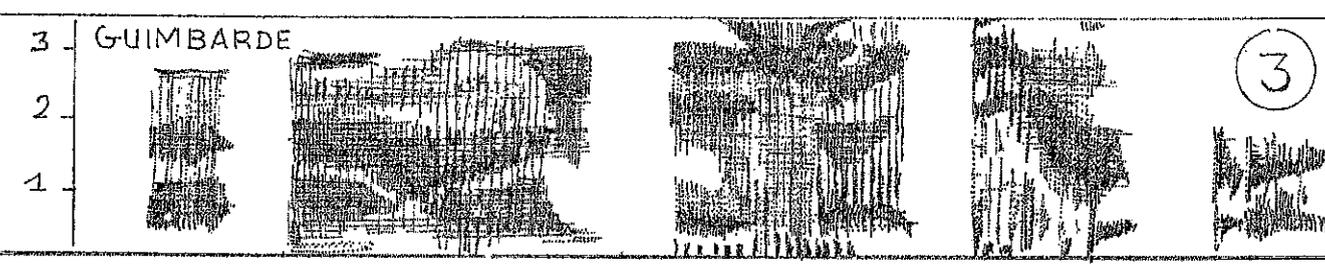
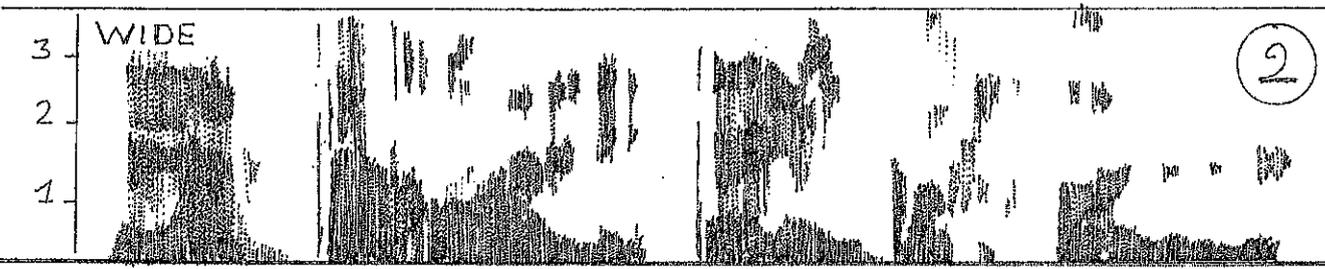


figure 7

A titre d'exemple, voici (fig.8) la série complète des sonagrammes intéressants pour une phrase donnée : " la parole est à l'Icophone ".

On y voit successivement :

..../



LA PAROL' EST A L'I_COPH.O_N.E

- a) le sonagramme normal de la phrase, entre 100 et 3000 Hz (analyse par bande de 45 Hz).
- b) le sonagramme en " wide " (bande de 300 Hz) : ici, du fait de la grande largeur de bande, le spectre de raies disparaît et les " formes ", dans le sens où nous l'entendons, apparaissent beaucoup plus clairement. Cependant les images sont trop floues : l'information esthétique (spectre de raies) y existe toujours quoique non visible et " brouille " les formes... On comprend cependant pourquoi les phonéticiens qui se sont servis du sonographe aient une prédilection pour cette représentation; nous trouvons là une confirmation de l'expérience de GEISSERT : le spectre de raies trouble la perception des formes en les quantifiant trop grossièrement. Cependant si nous possédons dans notre mémoire un stéréotype de ces formes, la corrélation est suffisante pour que nous retrouvions la forme dans le sonagramme, à condition de savoir où regarder...
- c) la même phrase à la guinbarde : la guinbarde, en supprimant le spectre de raies et le souffle donne des " résonances ", des formants, une image beaucoup plus nette que la voix normale : l'ordre de mouvement y apparaît de façon presque caricaturale. Mais malheureusement certains éléments phonétiques sont pratiquement irréalisables (on entend " icovone " au lieu de " icophone "). C'est pourquoi nous avons préféré utiliser la parole chuchotée.
- d) la parole chuchotée : Les formes y sont très nettement marquées le système phonatoire fonctionne normalement. Nous insistons sur le fait qu'il s'agit de formes floues, ce qui ne les empêche pas d'être significatives, originales et reconnaissables.
- e) la même phrase sur un magnétophone médiocre (dictaphone) Il ne reste presque plus rien en dessous de 1000 Hz et pourtant la parole est parfaitement intelligible : ce qui montre que l'essentiel du point de vue sémantique est contenu dans cette image.
- f) les formes sémantiques sous forme de " sténographié " comportant les éléments phonétiques reliés entre eux : c'est le modèle ^{utilisé} pour la synthèse.
- g) Ce qui reste du dessin précédent après synthèse à l'icophone. C'est peu de chose ; mais les formes ne sont toujours pas détruites : la parole reste intelligible; mais on est au bord de la destruction du message. De toutes façons, on vérifie, compte tenu des conditions d'anamorphose possibles, que ces formes se retrouvent dans tous les sonagrammes précédents; c'est évident puisqu'il s'agit précisément de l'information sémantique; la redondance n'est simplement pas la même d'un document à l'autre; en bref cette figure représente

...../

l'essentiel du message sémantique. Signalons que nous avons en construction un synthétiseur plus perfectionné qui nous permettra de reprendre et de développer ce travail.

VII. CONCLUSION

Le travail que nous avons réalisé sur la parole semble apporter un certain nombre d'éléments nouveaux en ce domaine. Nous avons donné une méthode simple pour extraire l'information sémantique, défini les grandes lignes d'une véritable Gestalttheorie de la parole et établi un dictionnaire des éléments phonétiques. Grâce à celui-ci, il est dorénavant possible de dessiner sans analyse préalable un mot, une phrase, un discours. Nous avons réalisé le synthétiseur qui nous a permis de vérifier nos hypothèses. Les résultats que nous avons acquis sont susceptibles de développements variés, dont certains sont en cours.

Il devient possible d'étudier la structure phonétique d'une langue, d'un style, d'un dialecte. Puisque l'intelligibilité de la parole est liée à des formes définissables, on peut étudier sous cet aspect les défauts de la parole et tenter de savoir en quoi, dans quelle mesure, et pourquoi les formes sont altérées. Malformations du système phonatoire et défaut de commande musculaire aboutissent forcément à des formes défectueuses.

Les télécommunications relatives aux messages parlés pourraient peut-être trouver ici des moyens d'économie : il est bien inutile d'encombrer les canaux avec de l'information parasitaire ou esthétique

La commande vocale de machines pourrait également y trouver son compte dans la mesure où on connaît la question de l'information sémantique.

De même il est évident que l'acoustique des " salles de parole " peut être abordée avec des idées nouvelles sur l'intelligibilité de la parole en fonction des caractéristiques de la salle. Nous avons d'ailleurs un travail en cours sur ce sujet.

On pourrait aussi reconsidérer la question de l'intelligibilité chez les chanteurs, etc....

En bref, si nous attachons une certaine importance aux résultats que nous avons obtenus, c'est surtout parce qu'il faut émerger des idées directrices nouvelles : on sait bien que des points de vue nouveaux permettent souvent de reprendre des questions qui semblaient insolubles. Nous espérons avoir apporté dans ce sens une contribution à l'étude de la parole, problème important, sur lequel travaillent d'arrache-pied de nombreux laboratoires dans le monde entier.

BIBLIOGRAPHIE

1. Mme METTAS (O) - Aperçu historique sur les appareils de synthèse de la parole.
Travaux de linguistique et de littérature
Ed. KLINCKSIECK STRASBOURG (1965)

2. LEIPP (E) - Le problème de la perception des signaux acoustiques par effet de contraste.
Annales Télécom. T. 20 n° 5-6 (1965) p. 103 - 118

3. LEIPP (E) - Un vocoder mécanique ; la guimbarde
Annales Télécom. T. 18 n° 5-6 (1963) p. 82 - 87

4. MOUNIER-KUHN et J.C. LAFON
De l'intérêt audio-phonologique de l'étude de la voix chuchotée au sonagraph
ACTA OTO-RHINO-LARYNGOLOGICA BELGICA Fasc.3 (1959)
p. 212 - 218.

Si l'on admet que la norme représente un certain état phonétique de la langue française, cet état semble être celui de la prose classique, du " bon français " enseigné à l'école, dont s'écartent notablement la poésie, le langage " parlé-écrit " de Queneau (dans " Zazie "), et les langages techniques. Bien sûr, entre ces divers langages existent toutes les nuances possibles. Mais tout se passe comme si, pour transmettre des idées par l'intermédiaire de la parole, dans un domaine non spécialisé (c'est-à-dire ne disposant pas d'un code spécial), un écrivain ou un orateur soit tenu de recourir au vocabulaire connu et compris de tous, et par là de respecter une norme phonétique représentative d'une langue.

IV - AUTRES APPLICATIONS ET CONCLUSIONS

Les tableaux de répartition des unités phonétiques nous conduisent à nous intéresser à de nombreux autres problèmes, en particulier au fonctionnement du système phonatoire : la fréquence d'occurrence d'une unité phonétique est-elle fonction de son mode d'émission ? On peut se demander par exemple, si l'existence du mot DE est responsable du premier rang qu'occupe l'unité DE dans la liste-type. Mais KE, KI, DU, Dè etc... sont aussi des mots courants, et n'en occupent pas moins des rangs fort médiocres; par contre l'unité AR, n'est pas un mot fréquent et on le trouve pourtant au 3^o rang. Ne serait-ce pas justement parce que l'unité DE présente une articulation " naturelle " en français, c'est-à-dire requiert un effort minimum pour un sujet parlant français, qu'elle coïncide avec un mot courant ? L'étude des distributions phonétiques dans d'autres langues et dans les premiers stades du langage enfantin devraient permettre de répondre à ces questions, que l'on ne peut aborder sans être armé d'hypothèses.

De même, quelles sont les unités phonétiques les plus rapidement altérées dans les diverses déformations du langage parlé, des patois et argots ? Quel rôle joue la vitesse d'élocution ? Comment se répartissent les durées des unités phonétiques autour de la moyenne de 50 ms en lecture rapide ? L'information transportée par une unité phonétique est-elle en relation directe avec sa fréquence d'occurrence comme le prévoit la théorie ? Quel est le mécanisme du décodage des formes phonétiques ? Voilà quelques unes des principales directions ouvertes à la recherche, à partir de l'existence d'une norme phonétique.

Celle-ci, comme la norme relative à l'emploi des mots (loi de Zipf), montre à quel point le langage est soumis au déterminisme statistique : quelles que soient les idées que nous voulons transmettre et les mots que nous choisissons il s'avère au bout d'un temps très court que nous employons toujours, sauf cas particulier, les mêmes proportions d'unités phonétiques.

Ainsi la notion de norme phonétique nous a permis de retrouver même sommairement, certains faits linguistiques : parenté ou caractérisation de textes. Peut-être même pourrions-nous grâce à elle identifier certains textes. Par suite elle justifie la considération des unités phonétiques comme éléments significatifs du langage, hypothèse que nous avons largement vérifiée sur le plan acoustique.