

CONFÉRENCES
DES
JOURNÉES D'ÉTUDES

FESTIVAL INTERNATIONAL du SON
HAUTE FIDÉLITÉ — STÉRÉOPHONIE
FACTURE INSTRUMENTALE

ÉDITIONS CHIRON, 40, rue de Seine - PARIS 6^e

L'INTELLIGIBILITÉ DE LA PAROLE DANS LE CHANT

par M. E. LEIPP

Maître de recherche au CNRS

Mlle M. CASTELLENGO

Laboratoire d'acoustique de la Faculté des Sciences de Paris

I. INTRODUCTION

La question de l'intelligibilité de la parole fait, depuis des décennies, l'objet de nombreuses recherches en divers laboratoires du monde entier, en raison de son incidence dans d'importants domaines technologiques et pratiques : télécommunications, messages vocaux dans le bruit, rééducation de la parole, etc. Le problème est actuellement loin d'être résolu de façon satisfaisante parce qu'il comporte de très nombreuses variables réagissant les unes sur les autres, et dont beaucoup sont encore mal connues. Comme il nous concerne directement au Laboratoire d'Acoustique de la Faculté des Sciences de Paris, nous avons fait et publié à ce sujet de nombreuses études (bib. 1, 2, 3, 4, 5, 6) qui nous permettent de tenter à présent une approche raisonnable du cas particulier de l'intelligibilité de la parole dans le chant.

Le chant est un phénomène acoustique résultant de la combinaison de deux types de signaux de nature différente fabriqués par l'appareil phonatoire humain : la parole et la « musique » des cordes vocales. Lorsque ce phénomène est associé à des sons instrumentaux qui l'accompagnent, le problème devient rapidement d'une complication incroyable. Pour s'en faire une idée, il est bon de se rappeler qu'en dernière analyse il s'agit de messages musicaux, problème que nous avons souvent abordé et dont un schéma de leur chaîne de communication va mettre en évidence les divers maillons dont l'étude sera notre objet ici (fig. 1).

Une telle chaîne suppose nécessairement un émetteur, un canal de communication et un récepteur qui, normalement, réagit au message.

Au début de la chaîne nous trouvons celui qui a conçu le message et l'a codé sous forme de partition : le compositeur. La partition qu'il écrit est un véritable programme de mouvements musculaires que doit exécuter le chaînon suivant, l'instrumentiste.

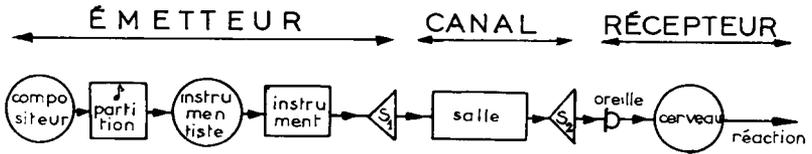


FIG. I. — *La chaîne de communication des messages musicaux.*

Elle met en évidence les divers problèmes qu'il faudra prendre en considération si on se propose d'étudier dans son ensemble le problème de l'intelligibilité de la parole.

Celui-ci, à son tour, fait fonctionner selon les règles la « machine à fabriquer les sons » que représente son instrument de musique.

Cette machine produit des signaux acoustiques, c'est-à-dire des variations de pressions aériennes compliquées, fugaces par nature, mais que nous savons maintenant saisir facilement avec le magnétophone et visualiser sous l'aspect de documents objectifs graphiques très significatifs.

Les signaux sont alors acheminés à travers un canal qui les déforme, les ampute de diverses manières et leur superpose un bruit de fond susceptible de les masquer plus ou moins.

Finalement le capteur, l'oreille, appréhende les vibrations aériennes et les transforme en impulsions électriques nerveuses ultérieurement décodées par le cerveau en fonction de références apprises et stockées en mémoire. Les références doivent être communes à l'émetteur et au récepteur, sinon le message ne pourrait avoir de signification, et la réaction désirée ne pourrait avoir lieu. Elles sont conventionnelles, comme le prouve l'existence de types variés de musique. Plongés dans un certain milieu sociologique, le compositeur et son auditeur sont soumis à l'apprentissage de la musique qui s'y fait. Apprentissage intuitif pour commencer, puis systématique avec de nombreuses règles d'association et d'exclusion. Les traces de cet apprentissage conditionnent à peu près exclusivement l'« homo musicus », souvent à son insu.

Passons à « l'instrument de musique » particulier dont il s'agit ici : l'appareil phonatoire, dont il convient de donner quelques détails structuraux et fonctionnels. Nous insisterons plus longuement sur la structure physique des sons émis par cet appareil dans le chant, car cela nous éclairera sur leurs déformations par le canal, déformations dont l'intelligibilité dépend essentiellement.

Lorsque nous aurons donné, enfin, quelques précisions sur la perception de la parole, nous aurons fait le tour du problème et nous comprendrons mieux à quelles difficultés redoutables les chanteurs ont à faire face, en particulier dans les œuvres lyriques.

II. L'APPAREIL PHONATOIRE ET SON FONCTIONNEMENT

STRUCTURE PHYSIQUE ET INFORMATIVE DE LA PAROLE ET DU CHANT

Puisque le chant est une combinatoire parole-musique, il est indispensable d'avoir quelques idées claires sur la « machine à parler » et les signaux qu'elle rayonne.

1. L'appareil phonatoire

C'est bien une machine, c'est-à-dire une association d'organes dont chacun possède un certain champ de liberté et dont les mouvements respectifs sont réglés par un programme, appris généralement dans l'enfance. Cette machine comporte deux grandes parties tout à fait distinctes, structurellement et fonctionnellement (fig. 2).

a) *Une série de résonateurs* réglables, dont 4 sont en série, et une, la cavité nasale est en parallèle. Ces cavités sont déformables volontairement dans une large mesure et de façon plus ou moins autonome. Chacune possède une fréquence propre de résonance, fonction de son volume et de la section de ses ouvertures. Les résonances affectent l'allure de bandes de fréquence assez larges : ce sont les « formants » de la parole. La figure 2b représente une simulation mécanique de l'appareil phonatoire dont les mouvements, y compris les ouvertures explosives de diverses « vannes » sont réglés par les cames d'un arbre. L'ensemble des cames représente le mouvement compliqué nécessaire pour articuler un mot.

b) *Les cordes vocales*. C'est un système d'anches doubles piloté par une musculature très élaborée. L'écartement des « lèvres » leur longueur et leur tonus sont réglés par des ordres impulsionnels venus des centres supérieurs.

Il est intéressant de préciser les éléments d'une typologie des signaux acoustiques possibles avec l'ensemble du système. Comme il s'agit de formes évolutives dans le temps, le diagramme fréquence-temps (sonagramme) nous fournira une représentation « parlante ». Les signaux rayonnés par le système phonatoire dépendent de la façon dont on l'excite. Deux cas sont à considérer, que nous allons étudier en détail.

2. Le système est excité par un écoulement d'air comprimé

Les cordes vocales sont au repos ; l'air comprimé est fourni par les poumons, et détermine des bruits d'écoulement classiques venant des frottements de l'air sur les parois.

Selon la disposition des organes, l'allure acoustique du signal correspond à une espèce de bruit blanc (fig. 3a) similaire à celui que

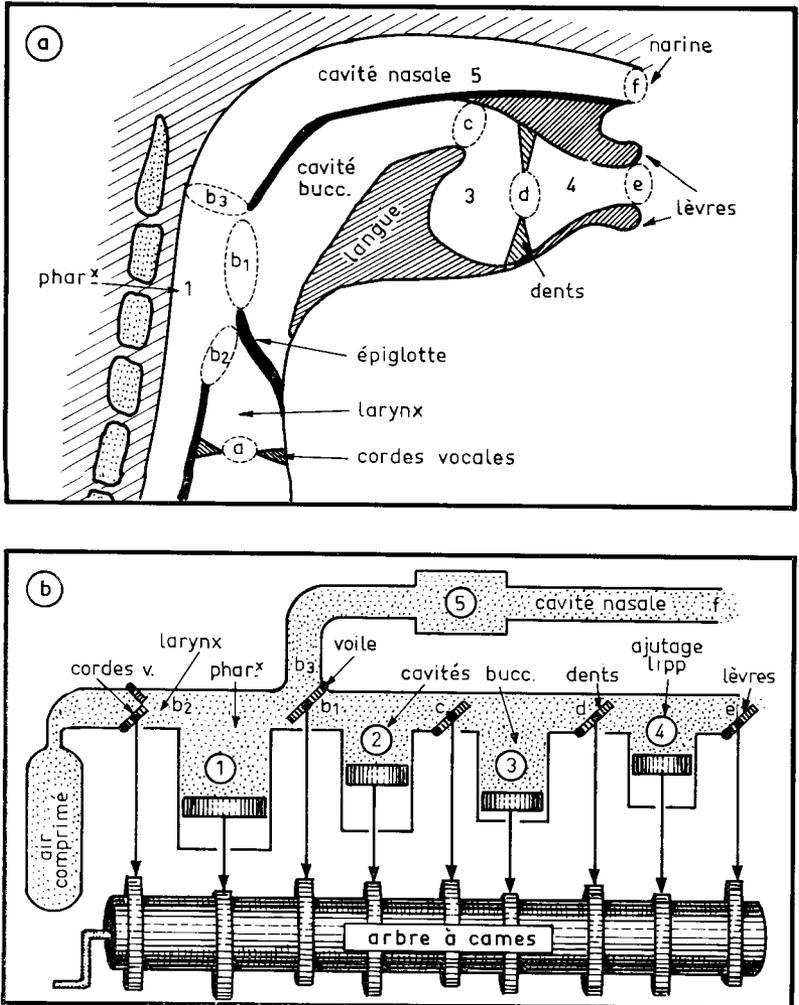


FIG. 2. — *L'appareil phonatoire humain et sa simulation mécanique.*

C'est une machine à faire des sons extrêmement élaborée comportant deux systèmes distincts : une série de résonateurs et des cordes vocales, qui contribuent à fabriquer les formes acoustiques de la parole et du chant.

l'on obtient en prononçant le son « f ». Pour d'autres dispositions on obtient le son « ch » (fig. 3b) ou le son « s » (fig. 3c). Dans tous ces cas, le frottement se fait au voisinage des ouvertures antérieures de la bouche (dents et lèvres). Comme les ouvertures sont alors très faibles, les résonances des cavités n'apparaissent guère.

Lorsque les ouvertures de sortie de l'air s'agrandissent, par

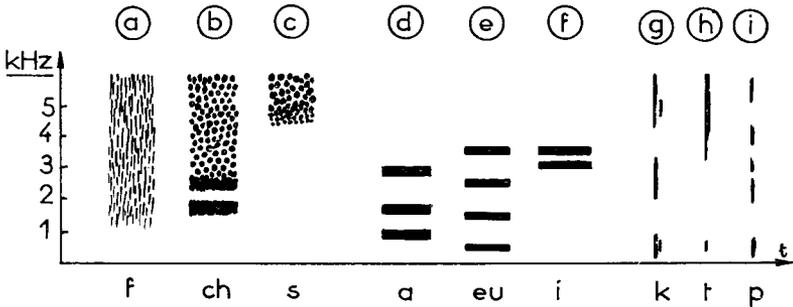


FIG. 3. — Typologie des sons élémentaires de l'appareil phonatoire excité par un écoulement aérien.

Ce sont des « bruits » dans le sens où on l'entend habituellement. Ils affectent l'allure de bruits blancs et colorés, de bandes de bruit ou d'impacts. Leur combinatoire et leurs enchaînements donnent la parole chuchotée.

contre, les bruits d'écoulement deviennent plus faibles, et chaque cavité détermine alors une bande de résonance appelée « formant ». Généralement plusieurs formants coexistent et leur intensité peut être différente ainsi que leur largeur de bande, selon les dispositions de l'appareil phonatoire. Ainsi un écoulement d'air dans les cavités disposées successivement comme pour prononcer « a », « eu », et « i » produit-il des signaux dont l'allure est donnée par les figures 3d, 3e et 3f.

Lorsque la langue provoque une explosion au niveau du palais, on obtient le son « k » (fig. 3g) ; au niveau des dents on a de même le son « t » (fig. 3h) ; si l'explosion est produite par les lèvres, on entend « p » (fig. 3i).

Tous ces bruits correspondent à un ordre de position des organes ; les sons obtenus représentent des « phonèmes ».

Mais la simple observation montre que lors de la parole normale, l'appareil phonatoire est continuellement en mouvement. La parole n'est pas une succession de phonèmes ; elle correspond non à un ordre de position des organes, mais à un ordre de mouvements. Les sonagrammes montrent clairement la différence : le mot « oui », n'est pas la somme de « ou » et de « i » (fig. 4a et b) ; le mot « chat » n'est pas la somme de « ch » et de « a » (fig. 4c et d) ; le mot « tôt » n'est pas la somme de « t » et de « o » (fig. 4e et f). Bref, un mot est une « forme » ; comprendre un mot c'est percevoir, reconnaître et décoder ces formes. Nous l'avons longuement vérifié en construisant un synthétiseur spécial, l'ICOPHONE, qui convertit des images optiques en « images » acoustiques, et nous en verrons les conséquences plus loin.

Les formes que dessine l'appareil phonatoire au moyen de bruits d'écoulement, à l'exclusion des cordes vocales, constituent la

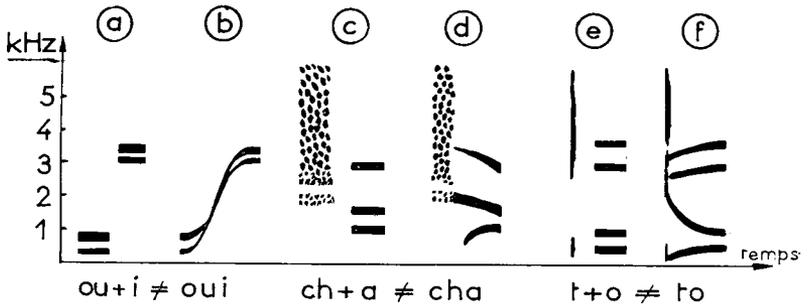


FIG. 4. — *Phonèmes et formes sémantiques vocales.*

Les phonèmes correspondent à un ordre de position de l'appareil vocal; les mots normalement articulés par contre correspondent à un ordre de mouvement. La différence est fondamentale: un mot n'est pas une séquence de phonèmes, mais une *forme* temporelle, une « *gestalt* ».

parole chuchotée; l'expérience montre que l'intelligibilité est totale: ce sont donc ces formes qui véhiculent toute l'information sémantique. Dès lors se pose la question de l'utilité et du but des cordes vocales en parole normale.

3. La parole normale : Le système phonatoire est excité par les cordes vocales

Des anches dissymétriques telles que les cordes vocales fabriquent nécessairement un signal en « *dents de scie* », riche en harmoniques de tous rangs (bib. 5). Ce signal peut être réglé en hauteur à volonté. Sur le sonagramme, un son grave se présente alors comme une série de raies parallèles, d'intensité décroissante vers l'aigu (fig. 5a); un son aigu est plus « *clair* » (fig. 5b).

Si on fait passer un tel spectre dans un système de cavités déterminant par exemple deux formants (fig. 5c), les harmoniques 2 et 6, qui tombent sur ces formants, sont amplifiés par résonance comme le montre la figure 5d.

Que se passe-t-il maintenant si les cavités sont en mouvement, cas de la parole normale? Un exemple simple va nous le montrer.

Soit le mot « *oui* », en parole chuchotée (fig. 5e) dont la forme sémantique est une espèce de double « *s* ». A chaque point de coïncidence de la forme avec l'un des harmoniques d'un spectre donné, 5a par exemple, il se produit une résonance, se traduisant sur le spectrogramme par un « *point-repère* » très net (fig. 5f). L'ensemble des points-repères nous redonne la forme en double « *s* », mais quantifiée assez finement, ce qui ne l'empêche absolument pas d'être reconnue. Le mot « *oui* » est donc parfaitement compris, mais comme le spectre de raies porte beaucoup plus loin

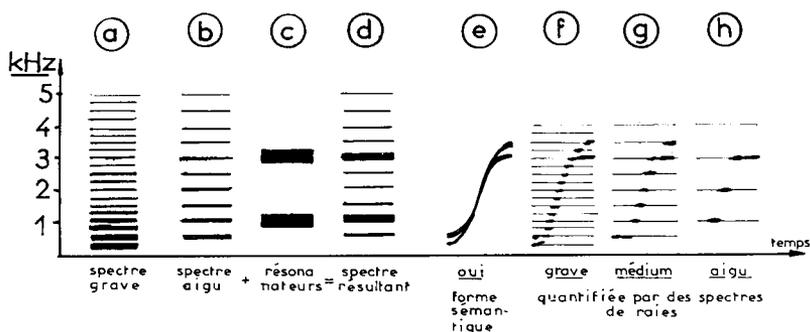


FIG. 5. — Spectres de raies vocaux et formes sémantiques quantifiées.

Les cordes vocales délivrent des spectres de raies plus ou moins serrées, colorés çà et là par les formants des cavités. Un mot normalement articulé à haute voix donne des formes d'autant plus finement quantifiées, c'est-à-dire reconnaissables, que le son est plus grave.

que les bruits d'écoulement de la parole chuchotée, le mot est compris à une distance bien plus grande : c'est l'intérêt du spectre de raies, et sa raison d'être dans la parole normale inventée par l'homme.

Si nous prenons un spectre de raies plus aigu, tel que celui de la figure 5b, qui est à l'octave du précédent, le même phénomène se produit. Mais, conséquence capitale, la forme est quantifiée plus grossièrement ; donc le nombre de points-repères diminue et la forme est partiellement détruite (dans le bas par exemple). Avec un spectre encore plus aigu, il ne reste que trois points-repères : la forme est difficile à retrouver, sauf si nous la connaissons au préalable. Dans ce dernier cas nous réussissons à la reconstituer mentalement et à la distinguer des autres mots comme « chat » et « tôt » sans ambiguïté.

La finesse de quantification diminue donc quand la voix monte ; mais corrélativement les cordes vocales deviennent plus petites et plus raides. Dans ces conditions, pour des raisons mécaniques, le son devient plus intense. Le chanteur aura donc à régler une difficile balance entre hauteur, intensité, timbre et intelligibilité, problèmes spécifiques du chant.

III. LES PROBLÈMES SPÉCIFIQUES DU CHANT DU POINT DE VUE DE L'INTELLIGIBILITÉ

Du point de vue de l'intelligibilité, le chant pose le même problème que la parole normale : il s'agit de la perception et de la reconnaissance des mêmes formes acoustiques. Mais les modu-

lations « musicales » des spectres de raies imposent à l'appareil phonatoire des dispositions et un tonus anormaux comparativement à la parole ordinaire, et les formes vocales seront anamorphosées en conséquence.

1. Anamorphoses et formes vocales chantées

Un raisonnement sur une forme graphique dessinée va nous permettre de poser le problème. Il n'y a strictement aucune différence entre les mécanismes de perception des formes graphiques et des formes acoustiques, et l'on retrouve ici toute une théorie de la forme (Gestalttheorie) élaborée au siècle dernier par Von Ehrenfels.

Prenons donc un « objet » bien familier, un lapin, dont nous possédons tous en mémoire un dessin stéréotypé (fig. 6a). Les deux coordonnées sont la hauteur en ordonnée, la largeur en abscisse, correspondant respectivement à la fréquence et au temps dans la représentation sonographique.

On vérifie instantanément qu'on peut rapetisser et déplacer la forme, l'écraser, l'allonger, etc... sans qu'elle cesse d'être parfaitement reconnaissable, et on ne la confondra jamais avec celle d'un sapin (fig. 6e).

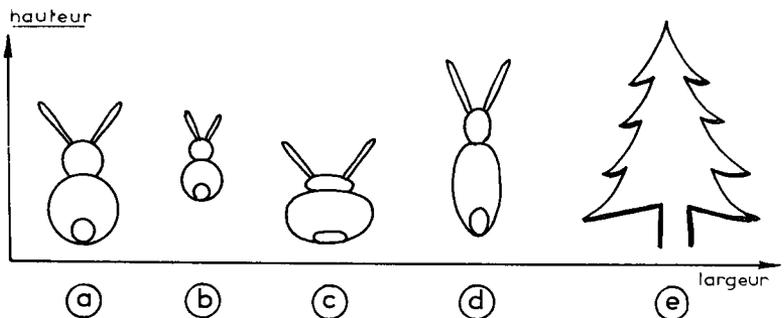


FIG. 6. — Formes et anamorphoses.

Une forme se reconnaît même si elle est fortement anamorphosée ; dans tous les cas elle ne se confondra pas avec une autre, différente.

Il en est exactement ainsi pour les formes vocales et un exemple réel va le montrer (fig. 7). On a pris la phrase « *mais oui messieurs* » dont par exemple la forme en « *v* » du début est particulièrement facile à repérer (fig. 7a). Articulons beaucoup plus nettement la même phrase. C'est bien toujours la même « super-forme » globale, malgré les différences d'allure : notre « *v* » est simplement un peu plus large (fig. 7b). On notera que la phrase 7a

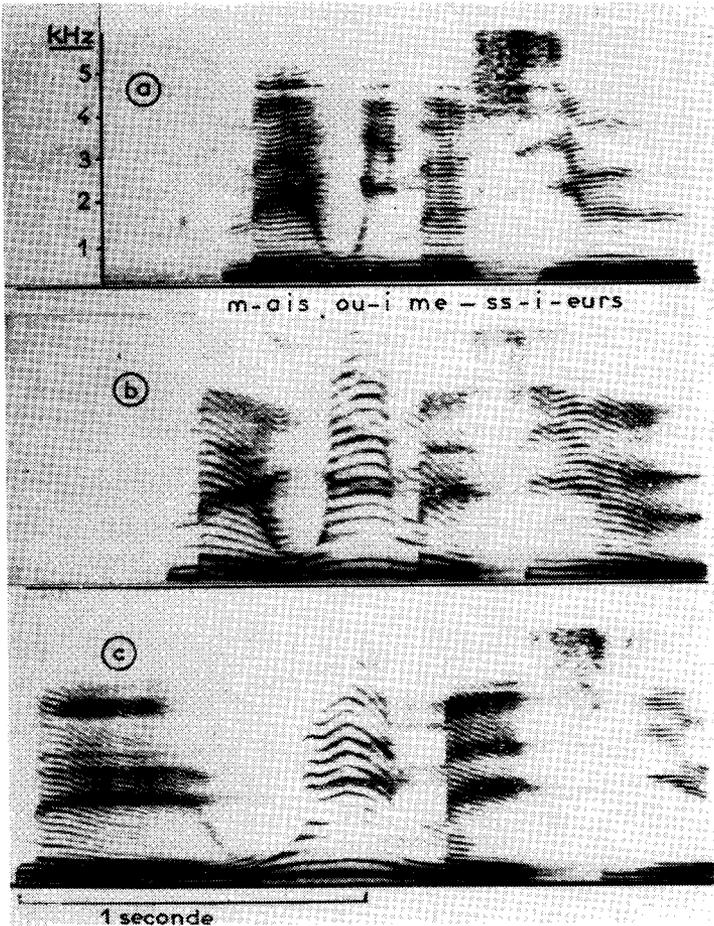


FIG. 7. — *Anamorphoses d'une phrase réelle.*

On articule la phrase « *Mais oui Messieurs* » de diverses manières et avec différentes intonations ; la « *superforme* » sémantique de la phrase est toujours reconnaissable : on comprend toujours les mots.

est prononcée en recto-tono, en voix monocorde : les raies ne changent pas d'écartement, en pratique. D'autre part, la phrase 7b indique des modifications de hauteur : l'intonation varie très largement, ce qui n'empêche pas la forme de rester reconnaissable, malgré une quantification locale plus grossière. Enfin la figure 7c représente la même phrase encore, mais prononcée très lentement. C'est bien toujours la même super-forme.

A partir de cet exemple, il devient possible de considérer le problème des formes acoustiques vocales chantées sous un aspect nouveau.

2. Le problème des formes acoustiques vocales chantées

En parole, les fluctuations de la hauteur se limitent à quelques tons, une quarte ou une quinte, parfois une octave dans les exclamations. En chant, on va beaucoup plus loin, et, au lieu de variations de hauteur quasi continues, on sélectionne pour articuler les mots, un certain nombre de « notes » discrètes, celles d'une gamme donnée. Comme on ne peut produire de spectres de raies que sur les voyelles, le chanteur commencera par s'entraîner à « tenir » des notes sur les diverses voyelles : ce sont les vocalises. Il chantera par exemple l'accord parfait sur « i » (fig. 8a), puis sur « a » (fig. 8b), etc. On notera que si la hauteur du son change, la « forme » déterminée par les formants reste identique pour chaque voyelle.

Ensuite, le chanteur apprendra à lier agréablement ces diverses voyelles entre elles (fig. 8c).

Finalement il sera conduit à prononcer des mots en vocalisant sur les voyelles. Il en résultera bien entendu une anamorphose particulière des formes vocales sémantiques, puisque certaines consonnes ne peuvent pas être modulées en durée. Mais, pour peu que l'on soit familiarisé avec le chant, on les reconnaîtra cependant, aisément, bref on comprendra les paroles.

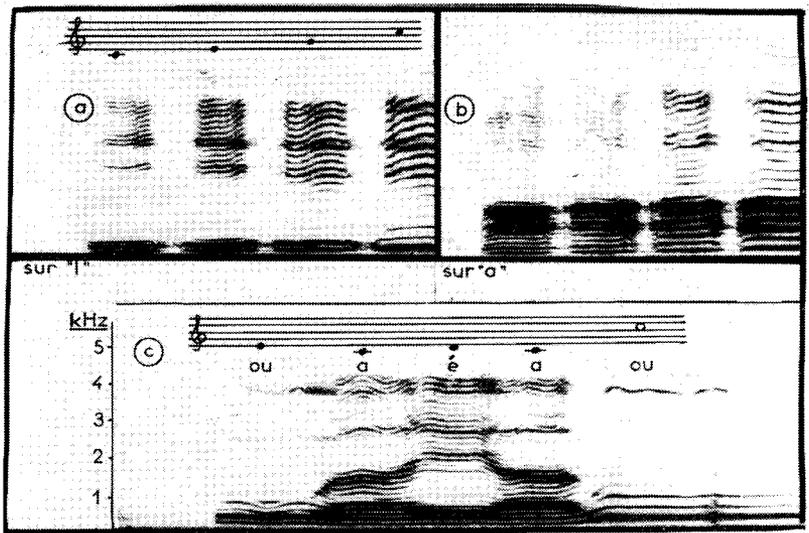


FIG. 8. — *Vocalises.*

Il s'agit de voyelles chantées sur des spectres de raies discrets correspondant à des notes de la gamme. Les enchaînements permettent de réaliser des sons musicaux bien ou mal articulés les uns aux autres : ils constituent une part importante de l'art de chanter et véhiculent une grande partie de l'information esthétique.

Un exemple réel (fig. 9) montre la forme acoustique d'une partie de phrase (*Miroir où se prend l'alouette*), extraite des Contes d'Hoffmann d'Offenbach. Cette phrase est d'abord articulée sur le rythme du chant, en parole chuchotée, ce qui met très clairement en évidence la superforme sémantique. Celle-ci est reconnaissable, et tout à fait différente de notre exemple précédent (*Mais oui Messieurs* !).

Cette même phrase a ensuite été chantée par une voix grave. La forme sémantique, finement quantifiée est parfaitement reconnaissable : nous comprenons tous les mots.

Puis, enfin, tout en bas, la phrase est reprise, une douzième au-dessus, par une voix de femme. La quantification est beaucoup plus grossière; la forme est donc beaucoup plus difficile à reconnaître, l'intelligibilité est moins bonne. Mais le découpage temporel reste chaque fois identique, et ce qui reste de la forme est suffisant

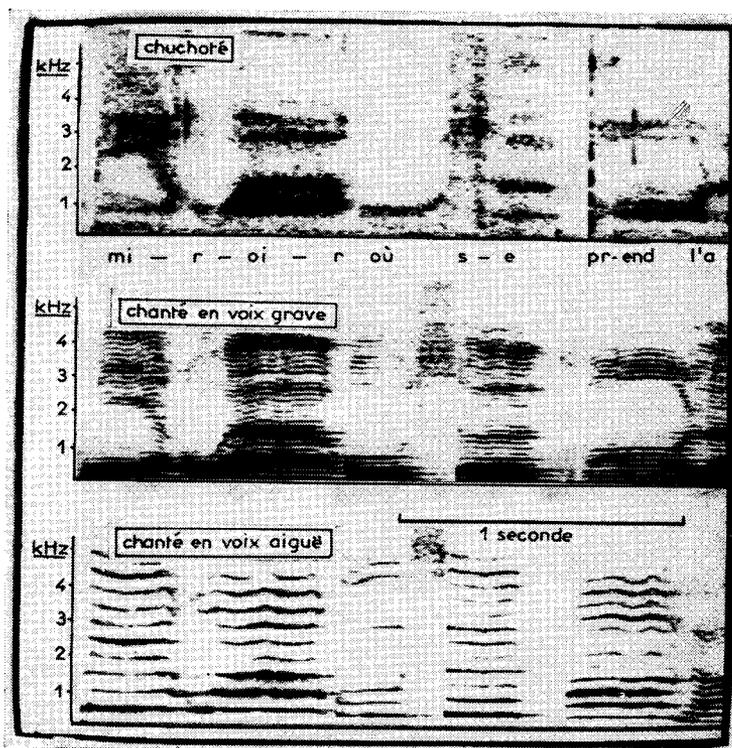


FIG. 9. — Spectrogrammes de chant normal.

On donne d'abord, en haut, la forme sémantique de la phrase, en parole chuchotée (*miroir où se prend l'a...*); puis la même phrase est chantée en voix grave : la forme sémantique est toujours perceptible. Avec une voie très aiguë, cette forme devient cependant moins évidente.

pour comprendre les mots grâce à divers indices (prononciation des « s » par exemple). De toute façon, l'intelligibilité est parfaite si nous connaissons par cœur les paroles, car alors les indices infimes suffisent pour qu'il n'y ait aucune ambiguïté. Nous retiendrons donc que, toutes choses égales, l'intelligibilité diminue avec la hauteur de la voix. Mais il va sans dire que de nombreux autres paramètres interviennent qu'il convient à présent de préciser.

IV. LES VARIABLES DE L'INTELLIGIBILITÉ DANS LE CHANT

Elles sont très nombreuses, et leurs interactions réciproques font du chant lyrique en particulier, un problème des plus difficiles qui soit. L'intelligibilité dépend en effet du chanteur lui-même, du contexte musical instrumental, des caractéristiques du canal (salle) et enfin de certaines particularités du récepteur.

1. Le chanteur

Ses dispositions anatomo-physiologiques conditionnent la netteté et la finesse des formes sémantiques au départ. L'entraînement lui permet bien entendu d'améliorer ses performances dans l'art de dessiner des formes sémantiques. En particulier, lorsqu'il va vers les registres aigus, il peut compenser la perte d'intelligibilité grâce à des artifices d'articulation bien connues des pédagogues du chant. Mais d'innombrables difficultés subsistent car lorsqu'on « force » la voix, soit pour monter plus haut, soit pour chanter plus fort, les dispositions de l'appareil phonatoire se trouvent modifiées et alors, de toute façon, les formes sémantiques émises sont moins évidentes pour un auditeur non familiarisé avec le chant. La nécessité de régler simultanément le timbre de la voix complique encore singulièrement la tâche du chanteur, et, aux anamorphoses complémentaires qui en découlent, il faut encore ajouter leur destruction partielle par le contexte instrumental normal.

2. Le contexte instrumental

Superposer à une forme acoustique vocale d'autres formes acoustiques, instrumentales, équivaut nécessairement à un masquage réciproque plus ou moins important. Nous avons longuement étudié ailleurs le problème du masquage d'une forme par un fond (bib. 7) et montré la nécessité d'établir de ce point de vue une typologie des formes acoustiques en cause. Voici donc l'allure de quelques sons musicaux classiques (fig. 10).

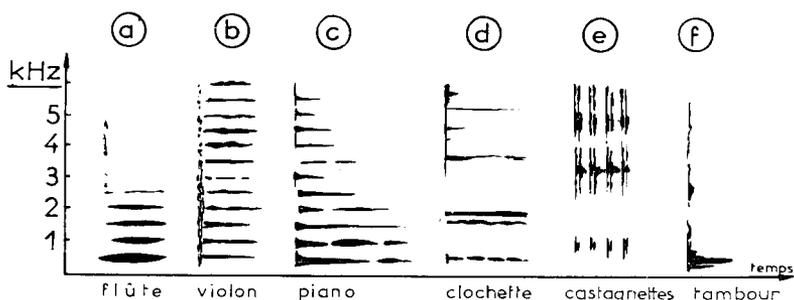


FIG. 10. — *Typologie schématique de sons musicaux.*

On y voit l'allure typique de sons de flûte, de violon, de piano, de clochettes, de castagnettes, de tambour. Ces « formes » musicales vont plus ou moins masquer les formes sémantiques de la parole.

On y voit un son de flûte, avec ses « fuseaux » caractéristiques (a) ; un son de violon, plus riche en harmoniques, avec le bruit de grattement de l'archet et l'attaque (b) ; un son de piano, quasi harmonique, avec ses partiels qui s'éteignent différemment (c) ; un son de clochette avec ses partiels qui battent plus ou moins ; des hachures verticales de castagnettes (e), et enfin un coup de tambour, hachure suivie d'une « résonance » grave (f).

Un exemple réel de superposition de tels sons avec du chant va montrer comment les formes vocales sont masquées (fig. 11).

En haut, on donne la phrase « *Mais oui Messieurs* » chantée sur une mélodie quelconque (do_3 mi_2 sol_2 do_3) (fig. 11a). On notera ici le vibrato de la voix, tout à fait classique dans notre musique occidentale et dont l'utilité va se justifier ici sans difficulté.

En effet, si nous jouons simultanément d'un instrument à raies harmoniques très riche (anches libres par exemple) en superposant au chant la même mélodie instrumentale, il est certain que le dessin vocal serait partiellement détruit par les raies horizontales de l'instrument (fig. 11b). Mais avec le vibrato, les sons musicaux et les sons vocaux se distinguent parfaitement. De plus la forme sémantique de la phrase se reconnaît sans difficulté (par exemple le double « v » de « *mais oui* ». S'il s'agit d'un instrument à sons variables, comme le violon, le hautbois, etc., qui permettent aussi le vibrato, il est évident que la voix peut être plus facilement couverte. Avec des instruments à sons fixes (genre orgue, régale, etc.), l'intelligibilité sera donc toujours meilleure, toutes choses égales bien entendu.

Reprenons la même phrase en agitant une clochette qui scande le rythme. Les trois raies spectrales de celle-ci ne masquent pratiquement en rien la forme sémantique (fig. 11c). Mais la clochette pourrait être une cloche, beaucoup plus grave et comportant beaucoup de raies plus ou moins fluctuantes et dont le son serait beau-

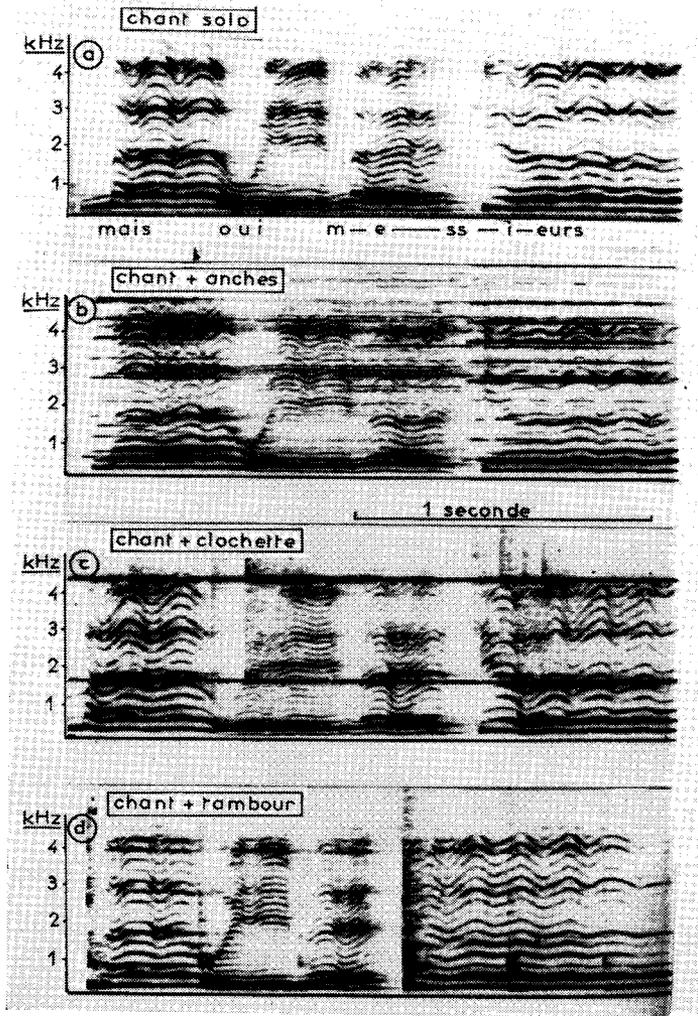


FIG. 11. — *Chant et sons musicaux réels.*

La même phrase, « *mais oui Messieurs* », chantée en solo puis partiellement détruite par la superposition de sons musicaux variés, montre que les formes sémantiques de la parole sont difficiles à détruire... Avec plusieurs instruments, le chanteur aura plus de mal à émerger !

coup plus intense ; la forme sémantique pourrait alors être complètement masquée et le message chanté serait inintelligible.

Battons enfin le rythme avec un tambour, dont on voit clairement les coups sous l'aspect de hachures verticales plus ou moins

larges. Le message vocal n'en est guère affecté. Il se pourrait que par hasard un coup de tambour tombe sur une plosive, par exemple « p » ou « t », le reste du mot n'étant pas détruit, ce qui reste de sa forme suffit pour que nous en fassions une reconstitution mentale, et cela d'autant plus facilement que nous connaissons le contexte du mot, ce qui nous permet de le prévoir très largement.

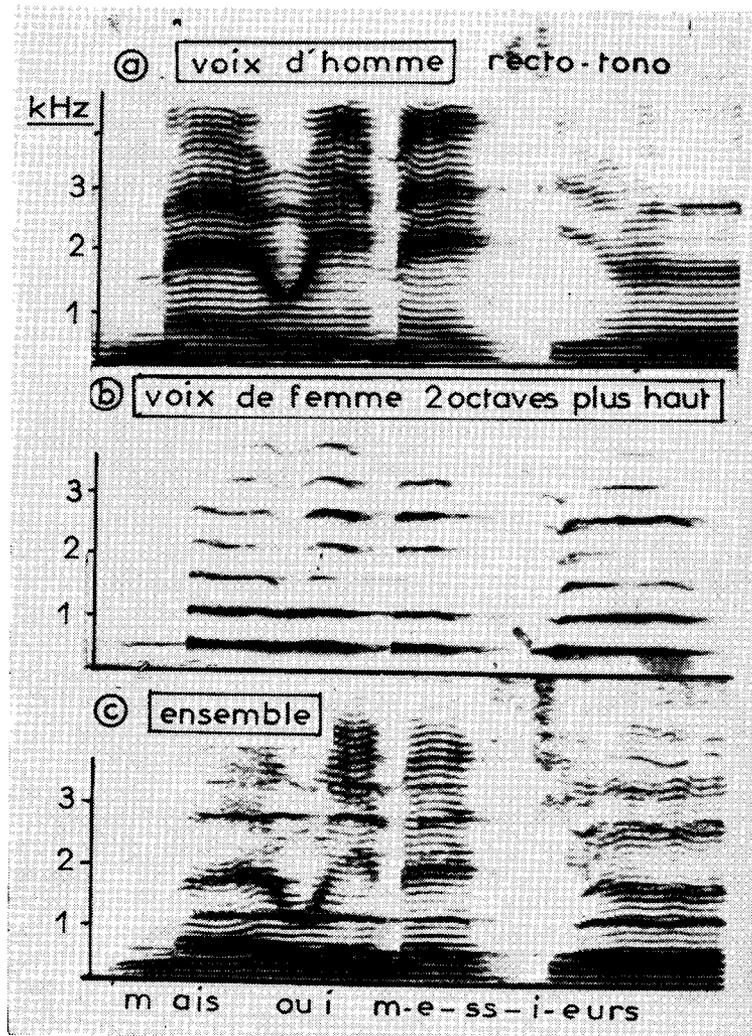


FIG. 12. — *Chant simultané d'homme et de femme.*

La forme sémantique est plus forte avec la voix masculine grave ; la voix féminine apporte des suppléments esthétiques de timbre.

On notera encore que si le coup de tambour est trop intense, il est susceptible d'« éblouir » notre oreille : le signal du mot prononcé existe bien, alors, mais nous ne le percevons pas, bien entendu.

Un cas particulier est à citer, qui se produit souvent dans les œuvres lyriques : un homme et une femme chantent la même mélodie et les mêmes paroles, à une ou deux octaves près. La figure 12 montre un exemple typique. On a chanté la même phrase « *mais oui Messieurs !* » en recto-tono. En 12a c'est une voix grave d'homme ; en 12b une voix de femme à la double octave. La figure 12c montre la superposition des deux voix. On notera que la superforme sémantique dessinée par la voix de femme est très approximative en raison de la hauteur de la voix. Dans le mélange, on comprend parfaitement les paroles, mais la voix masculine est déterminante de ce point de vue ; la voix féminine, par contre enrichit fortement le timbre. Cet effet est donc justifié. Il va sans dire que la moindre fausseté se traduirait par une non-coïncidence des raies : elle serait visible ici et audible de même.

On pourrait multiplier les exemples à l'infini, la combinatoire voix-instruments étant infinie dans notre musique. Mais il apparaît avec évidence que l'intelligibilité de la parole dans une œuvre lyrique dépend en premier lieu de l'habileté du compositeur à plaquer la musique adéquate sur le livret et à régler son orchestration de façon que l'émergence des formes sémantiques sur la musique soit toujours suffisante. Il y aurait toute une étude très intéressante à faire dans le cadre des classes de composition dans les conservatoires. Elle justifierait certainement ce qui est traditionnellement enseigné et mettrait sans doute en évidence l'habileté particulière de certains compositeurs célèbres, ceux en particulier qui, comme Berlioz et Wagner écrivaient leurs livrets eux-mêmes. Il va sans dire que l'habileté du chef d'orchestre à doser la musique peut jouer un rôle non moins déterminant dans l'intelligibilité, ainsi d'ailleurs que le canal, c'est-à-dire la salle d'audition, sur laquelle il convient de dire quelques mots.

3. Le canal

Nous montrons son importance capitale par un seul exemple réel (fig. 13). On sait qu'une salle peut avoir des échos, des « *retours d'images* » qui se superposent au son original. Lorsque des échos très rapprochés et indiscernables viennent de toutes les directions, on a une « réverbération » qui brouille plus ou moins les images acoustiques. Si les échos ne sont jamais recherchés en musique, il n'en est pas de même pour la réverbération, qui donne aux images acoustiques ce flou tant recherché, aussi en portrait photographique, parce qu'il efface de petits défauts et donne un « fondu » intéressant du point de vue esthétique.

Ici, nous touchons évidemment à une contradiction : pour avoir la parole la plus intelligible, la salle la plus sèche est préférée.

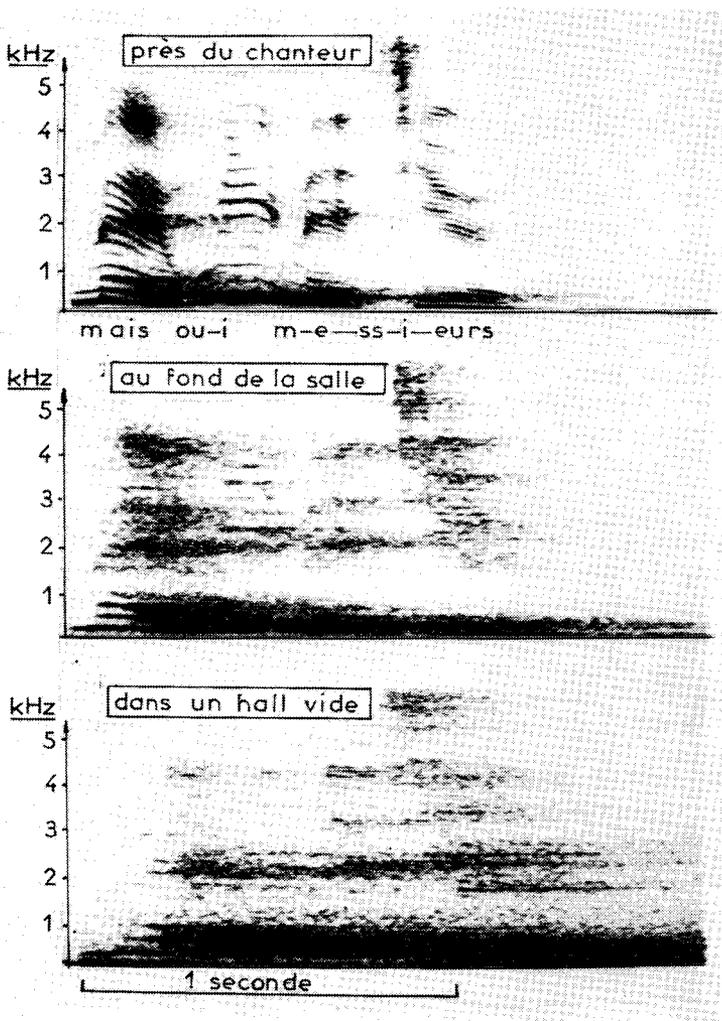


FIG. 13. — *Destruction de formes sémantiques par la salle.*

On voit à quel point une salle peut dégrader une forme sémantique en la brouillant et la filtrant. Le rôle de la salle est déterminant pour l'intelligibilité.

rable ; pour avoir de « beaux » sons musicaux, il faut une assez large réverbération. Or échos et réverbération détruisent plus ou moins les formes sémantiques et l'intelligibilité baisse alors au fur et à mesure que la qualité du son augmente. Une bonne salle est finalement celle qui réalise le compromis le plus heureux possible. L'ancienne salle du Conservatoire représente probablement un opti-

imum de ce point de vue, optimum atteint empiriquement et que des salles plus récentes sont loin d'égaliser parce qu'on n'a pas pris en considération lors de leur construction le problème de l'intelligibilité de la parole, pour la simple raison qu'il n'existe pas de doctrine cohérente et suffisante de ce point de vue.

La figure 13a nous montre ce qu'entend le chanteur lui-même dans un petit local vide, assez sonore, lorsqu'il chante la phrase « *Mais oui Messieurs* ». La forme sémantique est très nette, malgré quelques traînages sur les résonances de la salle, autour de 300 Hz par exemple, qui se voient très nettement vers la droite.

Au fond de la salle, l'enregistrement *simultané* montre à quel point la forme sémantique est déjà détruite. L'image est très claire au début, pendant un dixième de seconde environ ; puis tout se brouille. Pour nous, qui faisons l'expérience, la parole était parfaitement intelligible puisque nous la connaissions d'avance. La prévisibilité joue toujours un rôle capital en intelligibilité... Mais la même phrase chantée dans un hall est totalement inintelligible pour quiconque. Entre l'image du haut et celle du bas, on trouve évidemment toute la réalité des diverses salles existantes, dont la réputation statistique du point de vue intelligibilité est justifiable sans difficulté à l'aide de méthodes d'analyse comme celle que nous utilisons ici.

Il convient d'ajouter que l'intelligibilité d'un chanteur dépend encore de sa situation et de son orientation sur la scène, de la place de l'auditeur, des coulisses et décors, du taux d'occupation de la salle, etc. Finalement tout est encore compliqué du fait des particularités du système auditif de chaque auditeur.

4. Le récepteur humain

On sait que le capteur, l'oreille, varie du tout au tout avec les sujets. Un tel est sursensible aux fréquences aiguës là où un autre est sourd ; un tel dont l'oreille est amortie, physiquement, possède alors un capteur à grand pouvoir séparateur temporel. Des phénomènes de quelques millisecondes sont captés par lui, qui échappent à d'autres. Il perçoit un élément de forme sémantique vocale, là où son voisin n'entend positivement rien. Bref, selon les performances de l'oreille de chacun, l'intelligibilité de la parole d'une même œuvre peut sembler parfaite, douteuse ou mauvaise.

Mais ce n'est pas tout : dans l'énorme quantité d'information captée par l'auditeur lors de l'exécution d'une œuvre lyrique, il est impossible de tout appréhender. Le cerveau traite cette information en faisant un tri ; plus il est « puissant », plus il est susceptible de débrouiller des phénomènes compliqués plus ou moins simultanés. La tâche est difficile. Il s'agit d'une part de comprendre les mots, c'est-à-dire de repérer les formes sémantiques ; d'autre part,

pendant les vocalises, qui n'apportent rien du point de vue sémantique, nous explorons l'information esthétique supportée par les modulations variées des spectres de raies. Information sémantique et esthétique sont ainsi traitées séquentiellement, et leur « balance » judicieuse conditionne certainement la qualité des œuvres lyriques. Mais quoi que l'on fasse, l'intelligibilité de la parole dépend toujours, en dernière analyse des performances du récepteur, compte tenu de la prévisibilité des mots.

Il resterait bien d'autres éléments à analyser ; mais ceux que nous avons étudiés ici vont nous permettre de tirer quelques conclusions générales.

V. CONCLUSIONS

L'intelligibilité de la parole dans le chant pose des problèmes d'une complication incroyable, surtout lorsqu'il s'agit d'œuvres lyriques normalement chantées en salle, sans le concours de l'électro-acoustique... Ces problèmes, les chanteurs les maîtrisent généralement avec une habileté qui suppose de longues observations de la réalité chantée, un solide « métier », une intuition profonde. La science acoustique est actuellement fort loin de pouvoir les résoudre : nous serions mal inspirés de vouloir indiquer aux chanteurs ce qu'ils ont à faire ! Par contre, nous sommes parfaitement conscients que leur pratique est susceptible de nous éclairer sur mille problèmes qui nous concernent directement au laboratoire d'acoustique, et plus spécialement celui de l'intelligibilité de la parole, dans la mesure où nous acceptons de nous pencher avec beaucoup d'humilité sur ce qu'ils font. La présente étude ne prétend pas avoir épuisé le sujet, tout au plus espérons-nous en avoir cerné clairement certains aspects particuliers. De longues et patientes recherches restent à faire avant qu'il n'apparaisse en ce domaine une doctrine cohérente et suffisamment générale, qui pourrait alors, en retour, réagir efficacement sur les techniques d'enseignement des arts dramatiques et lyriques.

Nota : Nous développons actuellement des recherches sur les bases indiquées ici, grâce à la précieuse collaboration de Dominique PLESSIS et des musiciens de l'OPÉRA et de l'ORTF que nous ne pourrons jamais assez remercier pour leurs conseils et leur patience à nous exposer leurs problèmes.

BIBLIOGRAPHIE

- [1] LEIPP E. — Information sémantique et parole. Bulletin du Groupe d'Acoustique musicale de la Faculté des Sciences de Paris (GAM) n° 22. Paris 1966.
- [2] LEIPP E., CASTELLENGO M., LIENARD J.S., SAPALY J. — Structure Physique et contenu sémantique de la parole. Colloque GALF sur la parole (avril 1967), à paraître dans la Revue d'Acoustique. Paris 1969.
- [3] LEIPP E. — Le contenu informatif de la parole. 4^{es} conférences d'acoustique de Budapest. Comptes Rendus. Budapest 1967.
- [4] LEIPP E., CASTELLENGO M. et LIENARD J.S. — La synthèse de la parole à partir de diagrammes phonétiques. 6^e Congrès International d'Acoustique. Comptes rendus. Tokyo (1968).
- [5] LEIPP E. — Mécanique et acoustique de l'appareil phonatoire. Bulletin du GAM n° 32. Paris 1967.
- [6] LEIPP E. — Le problème de l'intelligibilité de la parole. Bulletin GAM n° 37. Paris 1968.
- [7] LEIPP E. — Le problème de la perception des signaux acoustiques par effet de contraste. Annales Télécomm. T. 18 n° 5-6 (1965), p. 103-118.

TABLE DES MATIÈRES

	Pages
Mécanismes physiologiques et perception auditive. (M. Legouix)	5
L'intelligibilité de la parole dans le chant. (M.E. Leipp, Mlle Castellengo)	15
Les microphones et la construction de la stéréophonie. (M. Condamines)	35
Une nouvelle méthode de jugement subjectif des haut-parleurs. (M. Burkowitz)	45
Mesures complémentaires sur les haut-parleurs et leurs enceintes. (M. Léon)	50
Sur les résonances et les réponses en acoustique des salles. (M. Bladier)	58
Un dispositif d'affaiblissement de bruit aux audiofréquences. (Dr Ray M. Dolby)	75
L'ionophone, pris comme source sonore d'accélération des molécules d'air pour les mesures acoustiques sur les instruments à vent. (M. Fransson)	91
Quelques problèmes posés par l'enregistrement de la musique traditionnelle des pays asiatiques. M. Tran-van-Khe).	109
Pouvoir des machines, pouvoirs de l'oreille. Introduction de l'électroacoustique dans la composition musicale contemporaine. (M. Reibel, M. Chiarucci)	118
Les machines à enseigner et la technologie de l'éducation. (M. Kirchberger)	125

Mécanismes physiologiques et perception auditive.

L'intelligibilité de la parole dans le chant.

Les microphones et la construction de la stéréophonie.

Une nouvelle méthode de jugement subjectif des haut-parleurs.

Mesures complémentaires sur les haut-parleurs et leurs enceintes.

Sur les résonances et les réponses en acoustique des salles.

Un dispositif d'affaiblissement de bruit aux audiofréquences.

L'ionophone pris comme source sonore d'accélération des molécules d'air pour les mesures acoustiques sur les instruments à vent.

Quelques problèmes posés par l'enregistrement de la musique traditionnelle des pays asiatiques.

Pouvoir des machines, pouvoirs de l'oreille. Introduction de l'électroacoustique dans la composition musicale contemporaine.

Les machines à enseigner et la technologie de l'éducation.



*organisées par le Syndicat des Industries
Electroniques de Reproduction
et d'Enregistrement (S.I.E.R.E.)*

16, rue de Presles, Paris 15^e,

sous le patronage de la Société

Française des Electroniciens

et des Radioélectriciens (S.F.E.R.), et du

Groupement des Acousticiens

de Langue Française (G.A.L.F.),

et du Comité Français d'Audiophonologie

au Palais d'Orsay, 7, quai A.-France, Paris 7^e



PARIS 1969