

Jacques CHAILLEY

EXPERIENCES DE CORRELATION  
ENTRE MUSIQUE ET PAROLE

Avril 1966

n° 19

G A M

BULLETIN DU GROUPE D'ACOUSTIQUE MUSICALE  
FACULTÉ DES SCIENCES — 8 RUE CUVIER — PARIS 5°

(Groupe d'Acoustique Musicale)  
 Laboratoire d'Acoustique  
 Faculté des Sciences  
 8, Rue Cuvier PARIS 5°

Adresse Postale :  
 9 Quai St Bernard PARIS 5°

BULLETIN N° 19

1°) REUNION DU 21 AVRIL 1966

Madame et Monsieur Zoltan KODALY nous ont fait le plaisir de nous rendre visite à cette occasion.

Monsieur GAUTHIER, Vice Doyen de la Faculté des Sciences nous avait honoré de sa présence.

Etaient présents :

M. le Professeur SIESTRUNCK, Président.  
 M. LEIPP, Secrétaire Général; Melle CASTELLENGO, Secrétaire.

Puis, par ordre d'arrivée :

M. FORET (ex Chef de la Garde Républicaine); M. SCHAEFFNER (Musée de l'Homme); Mme COSSART (Professeur de Musique); Mme de BOISSIEU (Professeur d'éducation musicale); Dr. HUET (Docteur en Médecine)  
 M. DUBUC (Musée des Arts et Traditions populaires); M. BERNARD (Maître de Conférences); Mme FULIN (CNRS). M. BLONDELET (Dir. Ets BUFFET CRAMPON); M. BATAISSIER (Secrétaire SIERE); M. RAVET (luthier); Mme NYEKI (Phonothèque Nationale); Melle PRADEL (Conservatoire); Dr POUBLAN (Médecin-biologiste) et Mme POUBLAN; Mme LEIPP; Mme HELFFER (Musée Guimet); M. BASCHET (ORTF); Mme METTAS (Phonéticienne); M. LE ROY (Professeur au Conservatoire); Melle WEBER et M. STOQUER (Institut de Musicologie); M. DUPARCQ (Dir. Revue Musicale) Melle Sylvie HUE et Melle RENAUDIE (Conservatoire); Dr. PERROT (Docteur ès Lettres); M. MARCIE (Hôpital Ste Anne); M. CHIARUCCI (ORTF); M. SCHWARZ (Musée de l'Homme); Mme BOREL MAISONNY (orthophoniste); Melle DINVILLE (orthophoniste); Mme KADI (Docteur en médecine); Mme et M. DORGEUILLE (Hôpital Ste Anne); Melle KLEIN (pianos KLEIN); Melle de MATONS (Etudiante); Melle Cl. MARCEL-DUBOIS (Directeur de Recherche CNRS).

Excusés : M. MAILLOT, M. TOURTE, M. FRANCOIS, Melle Odile VIVIER, M. CONDAMINES; M. ROUGET.

2°) Nous avons le plaisir de vous signaler la parution du " TRAITE DE LA FLUTE " de M. René LE ROY écrit en collaboration avec M. DORGEUILLE aux Editions Musicales Transatlantiques. En dehors de la technique et de la pédagogie de l'instrument qui sont la partie la plus importante, on trouve d'intéressants renseignements sur l'historique, l'acoustique et la fabrication de la flûte et de nombreux documents photographiques.

3°) M. TRAN VAN KHE revient des PHILIPPINES où il a pris part au Congrès international des Musiques Asiatiques; à son retour il a passé une dizaine de jours aux Indes et pris des contacts intéressants avec des musiciens. Grâce à lui et à M. RAO, joueur de vina, actuellement à Paris, nous allons pouvoir entreprendre une étude sur la question des échelles musicales aux Indes et nous espérons pouvoir faire une séance du GAM sur ce sujet, avant la fin de l'année.

---

EXPERIENCES DE CORRELATION ENTRE MUSIQUE ET PAROLE

Comme vous l'a rappelé la communication de M. RISSSET, la recherche de synthèse des sons, tant pour la musique que pour la parole, est très développée aux Etats-Unis. Durant mon séjour de 4 mois à l'Université de Californie, je n'ai pu obtenir que des renseignements assez vagues sur les travaux qui se poursuivent tant pour les autres Universités comme Columbia ou Princeton que dans les compagnies privées comme le Bell Telephone ou R.C.A. - Victor. De cette dernière firme seulement, j'ai pu obtenir copie, sans explication technique autre que ce que nous savons tous, d'un disque de musique synthétique dont avant d'aborder la relation de mes propres expériences à Santa Barbara, je vous ferai entendre quelques extraits.

Ex. 1 - R.C.A. - Victor, musique synthétique expérimentale.

Ce que vous venez d'entendre n'est plus, aujourd'hui, exceptionnel. Ce qui l'est davantage, c'est la brève séquence du même disque, sur laquelle les explications sont plus succinctes encore - on y devine même une volontaire réticence. Dans cet essai de chant synthétique féminin, vous observerez que le timbre de la voix humaine est, par endroits, reproduit avec un indice de perfection très satisfaisant, l'articulation des paroles par contre, est encore très loin du résultat à atteindre.

Ex. 2 - R.C.A. - Victor, chant synthétique.

J'en arrive maintenant à ma propre expérience. Celle-ci s'est trouvée provoquée par la rencontre faite, à Santa Barbara, l'un des 9 campus de l'Université de Californie, d'un collègue français linguiste, le professeur Pierre Delattre, établi depuis longtemps aux Etats-Unis, où il dispose d'un remarquable laboratoire d'étude de la parole et a bien voulu m'en accorder l'accès. Nous avons souvent et longtemps échangé nos expériences, et ce qui va suivre est le résultat de cette amicale et fructueuse collaboration.

L'appareil de synthèse de parole de Pierre Delattre a cette particularité d'utiliser exclusivement les sons harmoniques d'une fondamentale constante, de fréquence 120, soit approximativement un si bémol 2, dessiné sur une bande transparente défilant à vitesse réglable devant un faisceau lumineux que le dessin réfléchit et renvoie sur une cellule photo-électrique actionnant un générateur de fréquences. La parole est ainsi restituée : consonnes incluses, au moyen de seuls éléments de caractère musical, ce terme étant entendu dans son sens le plus étroit, à la fois consonantiel et intervallique. En voici un exemple :

Ex. 3 - Jacques Chailley synthétisé.

...../

Vous observerez que l'impression est celle d'un débit recto-  
tono sur le si b fondamentale, bien que ce son soit pratiquement ab-  
sent du dessin. Il en est ainsi de toutes les synthèses réalisées sur  
cet appareil.

Si je reprends mon nom ainsi dessiné et si j'écoute la même  
bande au ralenti, l'impression de parole s'affaiblira au bénéfice  
d'une perception musicale d'autant plus nette que le ralenti sera con-  
sidérable.

Ex. 4 - Jacques Chailley ralenti  
" " " très ralenti

Il semblerait donc possible, à partir d'une dictée musicale de  
ce genre, de reproduire les sons analysés ci-dessus sur un instrument  
de musique, un orgue par exemple, et de se rapprocher ainsi plus ou  
moins d'une sorte de synthèse musicale de la parole à partir d'un ins-  
trument de musique.

Il s'agit naturellement de tout autre chose que de la " vieille  
histoire " évoquée par M. LEIPP dans la discussion, consistant par  
MERSENNE à " faire parler " un tuyau à anche par des manipulations ana-  
logues à celles de la " Sourdine oua-oua " de nos trompettistes.

J'ai fait de nombreux essais. Après plusieurs échecs, je crois  
être arrivé à un résultat sinon probant, du moins encourageant, qui  
aurait sans doute été encore amélioré si mon travail n'avait été bruta-  
lement interrompu par des circonstances imprévisibles.

Le travail initial a porté sur les voyelles. Celles-ci nous  
enseignent les linguistes, sont produites par la combinaison, dans un  
rapport d'amplitudes défini, de divers sons fixes appelés formants, qui  
pour revêtir leur caractère spécifique, doivent se situer dans une zone  
de fréquence définie et relativement étroite - de 2 à 300 cycles pour  
les formants principaux. Plus exactement : la distinction linguistique  
entre deux voyelles est perçue par un minimum de deux formants. Le 3ème  
n'a pas grande fonction distinctive, car il est à la même fréquence  
pour toutes les voyelles (sauf i, où il passe environ de 2400 à 3000  
quand le 2ème formant est à 2500) pour une voix moyenne d'homme.

Pour que l'intelligibilité des voyelles reste acquise, il faut  
et il suffit que chaque formant soit situé à l'intérieur de sa zone  
avec les rapports d'amplitude voulus; mais pour que la voyelle ainsi  
émise devienne " timbrée ", ce que les linguistes appellent la parole "  
voisée "; il faut en outre que s'établissent entre ces formants des rap-  
ports de fréquence les situant comme harmoniques d'une même fondamen-  
tale, ce qui entraîne la perception de cette fondamentale ou de son  
octave comme hauteur de la voie entendue.

On distinguera donc deux notions différentes et complémentaires:  
l'une, l'intelligibilité, est affaire de hauteur absolue. L'autre, la  
musicalité, est indépendante de la hauteur absolue et ne dépend que de  
rapports d'intervalles.

...../

Nous avons essayé de reproduire à l'orgue les principales voyelles ainsi analysées. Les difficultés étaient considérables, la principale étant l'impossibilité de régler les rapports d'amplitude et d'éliminer les harmoniques parasites. Ici encore nos expériences ont été interrompues trop tôt pour pouvoir être jugées autrement que comme des indications provisoires. Voici, à titre d'exemple, les 5 voyelles A E I O U (= ou), présentées successivement et dans un ordre différent, d'abord à la machine de synthèse, puis, dans un autre ordre, à l'orgue. Pour éliminer toute intervention de suggestion, je ne vous annonce pas à l'avance l'ordre de leur présentation. Ce sera une expérience intéressante que de voir si malgré leur imperfection, vous identifiez l'ordre des présentations.

Ex. 6 - voyelles à la machine I E A O U (= ou)  
" à l'orgue A E I U O

(NB. - Cette expérience s'avère positive : l'ordre des voyelles est unanimement reconnu).

Si l'analyse des voyelles était jusqu'alors bien connue de l'ensemble des linguistes, celle des consonnes constitue, dit P. Delattre son apport personnel de découverte l'opposant encore à certains de ses collègues. On considérerait en effet, dit-il, le consonantisme comme un mélange de bruits inorganique, c'est-à-dire de partiels irréductibles au tableau des harmoniques, interférant avec les voyelles harmoniques. Sa théorie, démontrée par ses synthèses, est que la consonne est, comme la voyelle, réductible à la série harmonique, et résulte simplement d'une modification accélérée des formants, saisie dans l'instant de la transformation. Musicalement, une voyelle est un accord tenu ou se modifiant lentement; une consonne est une combinaison d'arpèges ou un contrepoint de lignes rapides ne quittant, comme les voyelles, l'échelle des harmoniques que par suite d'accidents dont nous parlerons tout à l'heure. Seuls sont inharmoniques les chuintantes ou sifflantes, s, ch, f, qui précèdent ou suivent les transitions, encore qu'on puisse parfaitement les synthétiser en utilisant exclusivement des accumulations d'harmoniques éloignés.

Une difficulté supplémentaire des synthèses de consonnes à l'instrument de musique - disons à l'orgue - est l'extrême rapidité des arpèges représentés.

Je n'ai pas osé m'attaquer à mon nom à cause de ses chuintantes et j'ai choisi la phrase " Arrêtez-le ". Le résultat est loin d'être parfait. Tel quel, il n'a pu être obtenu qu'en enregistrant une octave trop bas, au maximum de vélocité réalisable sur un clavier, et en doublant ensuite la vitesse de déroulement du magnétophone.

Ex. 6 - Arrêtez-le.

La phrase parlée se trouve ainsi ramenée à des éléments physiques communs avec ceux de la phrase musicale. Ceci rejoint l'expérience bien connue de la guimbarde : une même excitation sonore de l'instrument peut se traduire à volonté, selon la façon dont on modi-

fié la cavité buccale de résonance, soit par de la musique identifiable, soit par des phonèmes intelligibles.

Il était donc tentant, à partir de ces points de liaison, de chercher à préciser, en dialogue avec les linguistes, les rapports entre ces deux modes d'expression voisins, non pas sur le plan philosophique, où il n'est guère de frein aux imaginations les plus contradictoires, mais à partir de données expérimentales réelles et contrôlables.

Le premier point concerne le caractère musical même de la parole. On a vu que l'appareil de synthèse utilise exclusivement des sons harmoniques d'une même fondamentale, laquelle se trouve ainsi désignée - comme la note musicale identifiable de la voix parlée, ou éventuellement chantée même si elle est matériellement absente. Or il est bien certain que l'analyse au sonagramme de la voix parlée ou même chantée fait apparaître, à côté des harmoniques, des partiels inharmoniques dont il n'est pas tenu compte dans la synthèse envisagée. D'autres expériences auxquelles je n'ai pas assisté, ont établi, m'a dit M. Delattre, que si l'on mêle des partiels à une synthèse à base d'harmoniques ou si l'on déplace les formants de telle sorte qu'ils cessent de coïncider avec l'emplacement d'un harmonique tout en demeurant dans leur zone propre de fréquence, on ne porte pas atteinte à l'intelligibilité des phonèmes, mais on en trouble le timbre. Celui-ci devient, à mesure que les perturbations s'accroissent, de plus en plus voilé ou perturbé, pour finir, à l'extrême, par un chuchotement détimbré ou une voix éraillée. D'où l'hypothèse à vérifier que la présence de partiels dans un sonagramme de parole ne serait pas un élément de définition, mais d'altération. La parole la mieux timbrée ("voisée", disent les phonéticiens) serait celle qui contiendrait le plus d'harmoniques et le moins de partiels, et cette définition deviendrait l'idéal de la voix chantée. Il serait intéressant d'étudier, de ce point de vue, des sonagrammes de grands chanteurs. Par contre, cette notion n'affecterait pas l'intelligibilité du message sémantifié de la parole.

Cet idéal contient en lui-même une contradiction. On a vu que l'intelligibilité des phonèmes, et notamment des voyelles, était basée sur la position des formants sur une bande de fréquence limitée, en hauteur absolue, indépendamment des rapports d'intervalles. Lorsque par exemple la fondamentale monte d'une octave, soit 100 % en fréquence (rapport voix d'homme - voix de femme), la fréquence des formants monte seulement de 10 à 15 %. Le timbre, lui, est basé au contraire sur des rapports harmoniques d'intervalles, indépendants dans une certaine mesure, de la hauteur absolue. C'est pourquoi si un disque ou un magnétophone contenant de la parole tourne trop vite ou trop lentement, il n'y a pas seulement accélération et transposition, mais aussi distorsion; les voyelles notamment se trouvent modifiées, bien que le rapport entre leurs formants reste inchangé. D'où les difficultés des chanteurs à maintenir, sur une étendue qui dépasse la bande de fréquence en question, à la fois l'homogénéité de leur timbre et l'intelligibilité des paroles, surtout des voyelles. Bien souvent ils doivent sacrifier l'un à l'autre; les professeurs de chant le savent bien, et enseignent même par empirisme des déformations compensatrices de voyelles dans certains registres; ce n'est pas la moindre raison pour laquelle il est si difficile de comprendre ce que racontent sur scène les acteurs lyriques, et surtout les actrices ! Il pourrait, semble-t-il

y avoir là matière à des recherches spéciales dont la technique du chant pourrait retirer d'utiles enseignements.

Ce rapprochement de deux disciplines évoluent trop souvent, chacune de son côté peut également nous ouvrir des perspectives intéressantes sur la nature même du fait musical comparé à celui de la parole.

C'est naturellement entre parole chantée et parole parlée que les zones communes sont le plus étendues. On a vu déjà que moins un timbre parlé comporte de partiels inharmoniques, plus il se rapproche de la voix chantée. D'où viendrait dès lors la définition de la parole chantée ? D'abord bien sûr (Aristoxène l'avait déjà défini il y a 24 siècles), du caractère discontinu de l'échelle chantée face à l'échelle continue du parlé. Mais aussi d'un certain allongement du temps minimum alloué à chaque fondamentale. Celui nécessaire à la perception d'une voyelle est de l'ordre de 3 centisecondes. Celui d'un élément de consonne peut descendre jusqu'à 1 centiseconde. En musique, la durée minimum requise pour la perception d'une hauteur munie de sa signification individuelle ne descend guère au dessous de 20 à 25 cs., soit 1/5 à 1/4 de seconde (une double croche à 60 à la noire). Une note de musique, surtout instrumentale, peut certes être notablement plus courte, mais en ce cas la hauteur est plutôt perçue comme élément d'un ensemble à analyse globale (arpège, trait, etc...) que comme valeur individuelle. La parole chantée différencierait donc principalement de la parole parlée par le caractère non continu de l'échelle des fondamentales et par la durée de prolongation des voyelles sur une même fondamentale (un parlé recto-tono est déjà un élément de chant). La signification musicale spécifique du chant dépendrait ensuite de l'existence d'un rapport approprié entre ces diverses fondamentales, rapport auquel la parole parlée, en tant que telle, reste indifférente.

De la parole chantée, on passerait ensuite sans modification autre que celle des possibilités matérielles d'émission du son à la monodie instrumentale, telle qu'elle apparaît dans l'histoire avant l'entrée en scène de l'harmonie : simple décalque du sémantisme musical de la voix chantée, augmenté des possibilités spécifiques de l'instrument utilisé : vitesse accrue, ambitus plus ou moins étendu, ornements spécifiques, etc...

Des problèmes nouveaux, et qui n'ont été résolus en fait que par des siècles d'une lente et empirique expérimentation, se sont posés lors de la progressive apparition de l'harmonie, d'abord sous forme de simple consonance de contrepoint, puis plus tard - et sans doute dans la seule musique occidentale - comme élément déterminant du langage, réagissant à son tour sur la mélodie elle-même.

Mes travaux à Santa Barbara m'ont confirmé sur ce point les lignes directrices de mon Traité Historique d'Analyse musicale de 1950, tout en m'invitant à en préciser davantage certains aspects.

Le tableau des harmoniques, assoupli comme on sait par une tolérance définie aux approximations des échelles diverses, représente en quelque sorte un décalque en creux apte à recevoir, sous des conditions bien définies, les sons qui s'y adapteront pour les transposer



en perception globale consonante. Mais il ne s'agit nullement, comme on l'avait pensé un peu trop vite depuis Rameau, et surtout Helmholtz, d'une résonance matérielle assimilable à celle d'un corps sonore ou d'une corde vibrante. C'est pourquoi ce mot de "résonance", trop élargi par les musiciens pendant que les physiciens lui conservaient son sens limité, a faussé tant de dialogues entre les uns et les autres. Là où l'analyse d'un son donné nous livre (partiels exclus) une sélection d'harmoniques à peu près masqués par un son fortement prédominant, fondamentale ou non, le musicien superpose artificiellement une série de fondamentales d'intensité à peu près égale, génératrices chacune de leurs propres harmoniques dont il n'a cure. Le musicien élimine donc de sa zone de perception consciente tout ce qu'il utilise, volontairement ou non, à titre d'"harmonique de timbre".

Ce qui nous permet de dire, nonobstant toutes analyses de sonagramme qui ignorent cette distinction, que le musicien travaille en fait sur la prise de conscience d'un nombre d'harmoniques extrêmement réduit, absolument indépendant de celui qu'il utilise lui-même à titre d'harmonique de timbre, beaucoup plus réduit que lui, et plus réduit aussi que celui de la parole parlée.

Celle-ci en effet utilise, à titre de formants de timbre des fréquences de 200 à 3000 périodes environ, à titre de fondamentales des fréquences de 50 à 400 périodes environ. Ses formants représentent des rangs d'harmoniques de 1 à 40. La musique utilise théoriquement comme fondamentales des sons compris, selon les instruments, entre 32 et 2000 périodes, mais de fait rarement débordant de 65 à 1100 environ, ces fondamentales pouvant en certains cas être remplacées par des sons prédominants prélevés sur une série harmonique; elle ne tient compte qu'accessoirement des harmoniques que développent ces fondamentales (ou des harmoniques qui accompagnent le prédominant). Elle combine entre elles ces fondamentales (ou prédominants) comme s'ils étaient des sons purs, cherchant selon des principes spécifiques complexes à se rapprocher des rapports qu'ils prendraient entre eux s'ils étaient des harmoniques d'une même fondamentale.

Si l'on utilise artificiellement des harmoniques pour enrichir le timbre (ex. Mixtures d'orgue), on n'en tient pas compte en pratique dans l'acte compositionnel, sinon pour certaines "précautions" d'orchestration, d'ailleurs souvent négligées. Ces combinaisons ne dépassent pas, dans le rang des harmoniques, une limite variable selon les époques et les styles, mais dont le minimum est 4 et le grand maximum actuel 14. Je ne reviens pas sur la démonstration de ce fait, qui remplit tout au long mon Traité d'il y a 15 ans. Mais ce qui n'y était pas suffisamment développé - encore que cela y fût constamment sous-entendu - est le fait qu'il ne s'agit pas d'une superposition automatique et inintelligente, et que l'utilisation de ce phénomène de base, dont il ne se passe pas de jour que je ne voie quelque confirmation, suppose de la part de l'homme une opération personnelle exclusivement intuitive de sélection et d'adaptation sans laquelle il n'obtiendrait qu'un bruit inorganique.

Pour conclure notre propos de ce jour, il semblerait que nos expériences puissent aboutir à un certain nombre d'hypothèses de travail, à étudier en commun entre linguistes, physiciens et musiciens.

Parole et musique seraient, physiquement parlant, des phénomènes de même nature, basés sur l'utilisation instinctive mais contrôlée d'un mélange de partiels et d'une sélection d'harmoniques non perçus ni analysés un par un, mais utilisés en fonction de leur sonorité globale.

Dans la parole, les éléments nécessaires à l'intelligibilité des phénomènes que l'on entend charger d'un message sémantique peuvent être indifféremment harmoniques ou partiels, leur rôle étant déterminé par un certain ordre de fréquences attribuées à des formants définis en hauteur absolue. Mais la parole n'est " voisée " (les musiciens disent " timbrée ") que si ces formants se trouvent entre eux en rapports de sons harmoniques. En ce cas, une seule hauteur est perçue : celle de la fondamentale de ces harmoniques, qu'elle soit ou non exprimée. Il peut également y avoir hésitation entre la fondamentale et son octave, c'est-à-dire son harmonique le plus proche, d'où parfois hésitations sur l'octave perçue : cf. les pseudo-unissons d'hommes et de femmes. Par contre, l'élément musical de la parole, indépendant de son contenu sémantique et de la hauteur absolue, serait lié au caractère interval-lique de ces formants et à leur caractère de sons harmoniques. Ces deux éléments seraient donc indépendants l'un de l'autre.

La voix chantée serait susceptible soit de se confondre avec la parole parlée, soit de s'en éloigner selon le dosage respectif des deux éléments ci-dessus, mais aussi selon certains autres critères dont les plus importants seraient le temps de déroulement et l'échelle utilisée.

a) Considérée en fonction de la qualité du timbre, la voix chantée dépendrait, quel que soit le registre, de la répartition de ses composantes en harmoniques avec le moins de partiels possibles; mais en tant que véhicule d'un message sémantique, elle serait indépendante de cette notion et tributaire d'impératifs de tessiture parfois contradictoires avec ce qui précède, d'où les distorsions bien connues affectant l'intelligibilité des chanteurs; d'où aussi les truquages empiriques des méthodes de chant, qui pourraient peut-être tirer bénéfice d'études méthodiques dans ce secteur.

b) La perception de l'émission vocale en tant que hauteur étant indépendante de son timbre, pourvu seulement que celui-ci ait une majorité suffisante d'harmoniques sur les partiels, ne concerne que la fondamentale, ou éventuellement son succédané. Cette perception accuse d'autant plus son caractère musical que chaque fondamentale dispose de plus de temps pour être perçue comme telle. D'où le ralentissement de l'émission vocale chantée par rapport à l'émission vocale parlée.

c) Le caractère musical de la voix, indépendant de son sémantisme, apparaîtrait par les rapports de fréquences qui s'établissent entre les différentes fondamentales (ou jugées telles). Ces rapports seraient tributaires de lois spécifiques dont l'étude constituerait la base essentielle de la théorie de la musique. D'où l'aspect discontinu des échelles musicales plus propice à la mise en valeur de tels rapports opposé à l'aspect continu de l'échelle parlée, mais n'excluant pas l'hypothèse d'une filiation avec divers stades intermédiaires de transition, qu'étudie l'ethnomusicologie.



Un nouveau sémantisme spécifique de caractère proprement musical aurait pu naître ainsi, indépendant des associations d'idées convenues dont la parole se veut le véhicule, ; il permettra le développement progressif de la musique instrumentale, image de la vocale réduite au seul sémantisme musical spécifique (la vocalise y étant assimilée). La voix chantée, indépendante de tout accessoire harmonique, apparaîtrait ainsi comme un élément de transition auquel ferait suite dans l'histoire la monodie instrumentale également indépendante à l'origine de toute notion harmonique postérieure.

On n'a pas ici à poursuivre cette investigation plus avant : elle quitte notre propos de ce jour et nous introduit dans le domaine de phylologie musicale, dont nous ne faisons aujourd'hui qu'inventorier les prolégomènes.

De même l'apparition de l'harmonie verticale poserait de nouveaux problèmes, de nature différente, et qu'il importe de ne pas mélanger à l'aveuglette avec les précédents. On ne doit jamais perdre de vue en effet que le musicien manie et éventuellement superpose des sons complexes qu'il traite comme des sons simples, et les malentendus s'accumulent de ce chef dans son dialogue avec les physiciens.

Ce fait constitue sans doute l'un des principaux obstacles rencontrés dans mes tentatives de synthèse de la parole par la musique. Il n'en reste pas moins que, malgré l'imperfection des résultats très provisoires que j'ai cru pouvoir vous exposer, une voie de recherche reste ouverte, et c'est là la principale satisfaction que j'ai retirée de ces travaux trop tôt interrompus, et que j'espère pouvoir reprendre un jour.

DISCUSSION

M. LEIPP. L'analyse et la synthèse de la parole sont une de nos préoccupations au laboratoire; les expériences de M. CHAILLEY montrant que l'on peut fabriquer une parole intelligible à partir de sons musicaux nous intéressent donc vivement. Un certain nombre d'observations que nous avons faites semblent montrer cependant que les spectres de raies ne jouent aucun rôle dans l'information sémantique du discours; nous l'avons montré naguère à l'aide de la guimbarde, instrument de musique primitif fabriquant un spectre de raies fixe et qui permet de réaliser une parole synthétique parfaitement intelligible (\*). Mais la contradiction n'est qu'apparente : on sait qu'un son musical très bref est perçu par notre oreille comme un " clic ", c'est-à-dire un petit bruit; en additionnant des petits clics on peut réaliser pour l'oreille n'importe quel signal aléatoire, donc les chuintantes et les sifflantes : c'est une question de vitesse de défilement des phénomènes.

Nous ajoutons que la synthèse de la parole préoccupe les chercheurs depuis fort longtemps, et que Mme METTAS prépare actuellement une thèse sur les méthodes historiques utilisées. De notre côté, signalons le chapitre de l'HARMONIE UNIVERSELLE, où MERSENNE nous dit " comment il faut construire les jeux d'orgue pour prononcer les voyelles, les consonnes, les syllabes et les dictiones. " .... Ce n'est pas tout à fait l'idée directrice de M. CHAILLEY, mais le problème était posé.

Je voudrais savoir si l'intelligibilité de la parole synthétisée par M. DELATTRE est suffisante.

M. CHAILLEY. M. DELATTRE nous a fait entendre un poème d'APOLLINAIRE où l'on ne perdait pas une syllabe...

Mme BOREL MAISONNY. Il faut être prudent : l'expérience nous a montré qu'à partir du moment où l'on sait ce qui va être dit, la perception devient facile à partir d'éléments informes.... D'autre part, il faut considérer que les consonnes présentent deux aspects différents selon qu'il y a de la voix ou non; avec des consonnes du type "s" le caractère vocalique manque totalement; on ne peut donc les synthétiser avec des sons musicaux.

M. CHAILLEY. C'est exact; certaines consonnes ne peuvent être reproduites avec des harmoniques identifiables en tant que tels. D'autre part, les occlusives se traduisent sur les graphiques par des silences préparés par la voyelle précédente et annonçant la voyelle suivante; un exemple typique est le "T" dans la phrase "arrêtez-le!".

..../

---

(\*) Un vocoder mécanique : la guimbarde Annales des Télécommunications Tome 18 (5-6) 1963.

Mme BOREL MAISONNY. C'est la théorie du locus, dont on a beaucoup parlé; mais elle est insuffisante.

M. CHAILLEY. Les expériences que j'ai faites avec l'orgue montrent en tout cas qu'il y a une grosse perte d'intelligibilité :

1°) parce que les harmoniques des instruments utilisés pour mes expériences n'étaient pas justes.

2°) parce qu'on ne peut pas modifier les amplitudes des composantes à l'orgue.

3°) Parce que nous utilisons des sons musicaux qui ne sont pas purs et qui surajoutent leurs propres harmoniques. Mais on doit pouvoir y arriver : c'est une question de patience.

Mme NYEKI. N'avez-vous pas essayé d'autres instruments, la flûte par exemple ?

M. CHAILLEY. Non! celle-ci ne descendrait pas assez bas et on aurait moins de précision.

M. GAUTHIER. Par rapport aux expériences antérieures celles de M. CHAILLEY apportent un élément nouveau et important : la possibilité de ralentir des phénomènes acoustiques en conservant leur hauteur. Elles montrent aussi que pour l'oreille certains phénomènes temporels sont " avalés " par la vitesse, alors que d'autres, en succession rapide fusionnent en un phénomène unique dont la structure n'apparaît qu'à partir d'une certaine vitesse.

Nous remercions M. CHAILLEY de nous avoir apporté ses résultats et d'avoir su nous les exposer de façon aussi intéressante.

