

Les sciences de la forme aujourd'hui

Depuis Platon et Aristote, l'étude des formes et de leur croissance – ou morphogenèse – s'est lentement dégagée de ses présupposés magiques et de son contenu esthétique. Grâce aux travaux de quelques pionniers, dont D'Arcy Thompson ou Alan Turing, et aux avancées des mathématiques, de la cristallographie et de l'embryologie, la morphogenèse est devenue un thème de réflexion majeur traversant tous les domaines du savoir.

Mathématiciens, chimistes, biologistes, mais aussi musiciens, architectes et linguistes exposent ici leurs conceptions de la forme et tracent les grandes lignes d'une nouvelle approche – plus qualitative et plus globale – de la recherche scientifique.

Y. Bouligand, M. Castellengo, G. Cohen-Tannoudji,
Y. Couder, F. Dagognet, J. Dhombres, J.-P. Luminet,
A. Prochiantz, A. de Ricqlès, J.-P. Saunier,
R. Thom, C. Vidal, P. Virilio, M. Yaguello,
interrogés par Émile Noël

Une série d'émissions diffusées sur France-Culture
dans le cadre des « Perspectives scientifiques »



Seuil, 27 r. Jacob, Paris 6
ISBN 2.02.022649.9 / Imp. en France 10.94

cat. D

Y. Bouligand, M. Castellengo, G. Cohen-Tannoudji,
Y. Couder, F. Dagognet, J. Dhombres, J.-P. Luminet,
A. Prochiantz, A. de Ricqlès, J.-P. Saunier,
R. Thom, C. Vidal, P. Virilio, M. Yaguello
interrogés par Émile Noël

Les sciences de la forme aujourd'hui



Les sciences de la forme aujourd'hui

6102

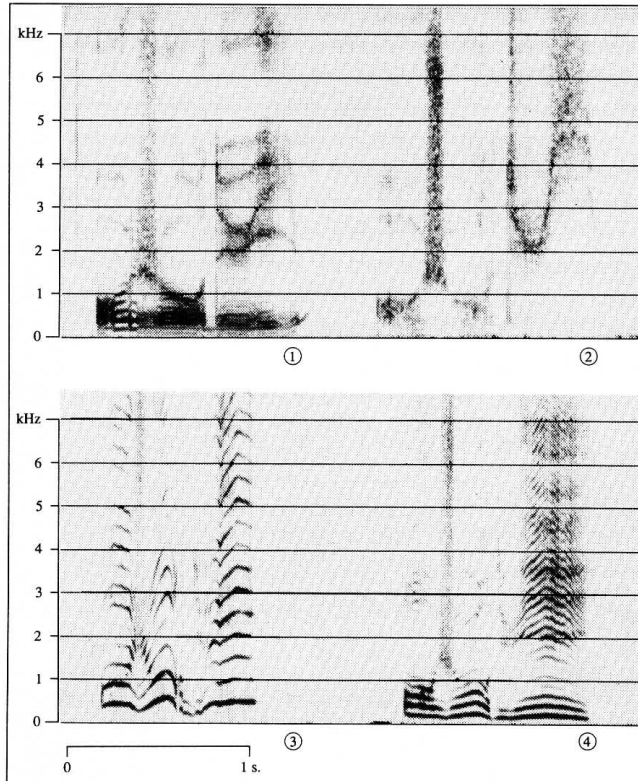
Inédit Sciences

10. Les formes sonores

Entretien avec Michèle Castellengo, directeur de recherche au CNRS, directeur du Laboratoire d'acoustique musicale de l'université Paris-VI.

ÉMILE NOËL : Dans le domaine du son, la notion de forme va de soi. On parle couramment de forme musicale. Il est vrai que cela vise plutôt les formes esthétiques et les genres musicaux. Mais cela a-t-il un sens de supposer des formes sonores élémentaires comme on parle de particules élémentaires ?

MICHÈLE CASTELLENGO : Peut-être. Cela peut avoir un sens si l'on pense à la manière dont nous identifions les différentes sources. Par exemple, actuellement, nous sommes en train de réfléchir au laboratoire sur des actions comme frotter, souffler, pincer, frapper. Si nous n'avons pas reconnu une source sonore, nous pouvons tout de même dire : c'était un frottement, un craquement, un raclement, un sifflement. Et je crois que le simple fait d'avoir reconnu, à l'écoute du son, la manière dont il a été produit, c'est-à-dire le geste qui est cause de la vibration, est une reconnaissance de formes élémentaires. Il existe dans le signal sonore des éléments que nous reconnaissons et qui nous



Le même mot *aujourd'hui* prononcé par quatre locuteurs différents (voir texte). La forme acoustique, caractéristique de ce mot, est reconnaissable à travers les diversités des voix. Diagrammes temps/fréquence.

permettent d'identifier tout de suite le son et, au besoin, de le nommer.

Parler de formes sonores, en l'occurrence, c'est parler de formes sonores entendues et reconnues.

Bien sûr. En tant qu'être vivant, nous décodons notre environnement en allant y puiser les formes qui nous intéressent. Cette façon de procéder, très efficace, suppose que nous ayons auparavant élaboré plusieurs répertoires de formes mises en mémoire. Et reconnaître la source qui produit un son, c'est bien comparer cette nouvelle forme sonore perçue à une autre semblable, mémorisée, produite par la même source.

Mais alors, cette forme élémentaire est déjà, en quelque sorte, une forme globale.

C'est la notion même de forme qui est globale, au sens où c'est un objet défini par des rapports, un tout qui n'est pas la somme des parties, que l'on appréhende dans son ensemble au travers de la variabilité survenant à chaque présentation.

Objectivement, un son ou un groupe de sons, ce sont des vibrations. Comment analysez-vous les corps sonores ? Comment les reconnaissez-vous ? Les vibrations donnent-elles des formes ?

Il y a plusieurs manières de procéder, et la représentation la plus convenable dépend de l'objectif de l'étude. Dans le cadre d'une étude physique, nous devons avoir accès à un certain nombre de grandeurs très précises des constituants du son, et il existe quantité d'outils numériques qui nous

donnent satisfaction. Mais si nous voulons représenter l'information, donc les données mesurées, sous une forme interprétable par un être humain, nous allons privilégier la mise en relation de l'évolution d'un certain nombre de données sous une forme visuelle facilement appréhensible.

En ce qui concerne le comportement d'identification et de reconnaissance des sources que je vous ai cité tout à l'heure, nous utiliserons des représentations qui montrent l'évolution de la fréquence et de l'intensité en fonction du temps, et qui produisent des formes visuelles tout à fait analogues aux formes sonores mémorisées. Mais cette analogie, qui a l'air de bien fonctionner, et qui est heuristique, comporte un petit inconvénient : traitant simultanément plusieurs dimensions, elle n'a plus la même précision sur chacune d'elles.

La vibration d'une corde de violon ou de la table d'un instrument quelconque peut donc être représentée par un dessin dont la forme va se retrouver à peu près régulièrement pour toute semblable animation de la table ou de la corde.

Oui. En tout cas, pour revenir à l'exemple que je vous donnais au début, si cette corde a été pincée ou si elle est frottée avec un archet ou tout autre mode d'excitation, l'analyse nous donnera les indices qui nous permettront de faire l'association entre la reconnaissance à l'oreille et le tracé. Or ces indices sont temporels. Il ne faut pas oublier que, si l'on parle de formes en acoustique, ces formes se déroulent dans le temps : elles ont un début et une fin. Notre traitement de l'information sonore est hautement attentif à ce qui évolue dans le temps. Un son qui ne change pas – un son complètement périodique, une sinusoïde, ou même un son très complexe mais qui ne varie pas

– est non seulement fatigant mais, peu à peu, il sera oublié. En 1636, le père Mersenne, dans *l'Harmonie universelle*, dit en substance : « Oyez un son de flûte éternellement, il vous ennuiera et, à la fin, il vous fera mal à la tête. » C'est bien l'évolution dans le temps qui définit la forme acoustique perçue.

Les formes dont vous parlez ne sont pas des sinusoïdes, des représentations graphiques, mais des formes qui ressemblent à des dessins.

Effectivement. Les diverses représentations visuelles du son n'ont pas toutes la même prégnance perceptive.

Prenons, par exemple, le tracé oscillographique, qui est dans toutes les mémoires et que l'on voit souvent dans les manuels d'acoustique. Il n'est pas directement fiable à la perception, et ce pour deux raisons. Tout d'abord l'échelle temporelle de représentation, la période, est extrêmement brève. On ne peut guère faire de relation entre quelques millisecondes de signal et une perception. D'autre part, la forme visuelle du signal sonore amplitude/temps, si précieuse pour l'étude physique du son, est très sensible aux variations de phase des différentes composantes, alors que l'oreille ne l'est pas. Autrement dit, deux sons donnant des sensations sonores identiques pourront avoir des représentations visuelles très différentes. Si bien que l'on en revient aux formes les plus interprétables en termes de perception. Si l'on se rapporte à la manière dont nous décodons l'environnement sonore, qu'il s'agisse de bruits, de sons musicaux ou de paroles, nous pouvons constater que nous avons tous une compétence extraordinaire qui se développe peut-être déjà avant la naissance et qui, très vite, devient très efficace chez les bébés. Sur la base d'une détection fine des changements du signal, nous structurons

des formes dans la dimension temporelle des événements sonores.

En parlant de formes qui se différencient des diagrammes ou des spectres vibratoires, je pensais à des formes presque figuratives, qui peuvent, en tout cas, faire penser à une figuration. Me venait à l'esprit que ce que vous étudiez, vous ne pouvez guère vous le représenter sérieusement que de façon visuelle ?

C'est bien à contrecœur que je suis obligée de vous l'avouer. Dans le domaine acoustique, on est tributaire de la représentation visuelle pour faire une étude objective des phénomènes. Et c'est bien ce qui explique aussi le grand retard des études de psycho-acoustique par rapport à celles de la perception visuelle : il a fallu attendre de pouvoir enregistrer les sons et d'en avoir des représentations visuelles objectives. C'est grâce à cette représentation visuelle qui est là, sur la table, entre nous deux, que nous allons pouvoir parler des sons et les décrire. Autrement, nous n'aurions accès qu'au souvenir du son, à la trace, restée en mémoire, de ce que nous avons entendu. Le deuxième progrès important – à part la représentation visuelle à laquelle se sont efforcés de parvenir tous les chercheurs – est la possibilité d'avoir une représentation visuelle de l'évolution temporelle des phénomènes sonores. L'analyse acoustique des sons évolutifs, le traitement des transitoires, comme on les désigne en acoustique, ne se fait correctement que depuis une cinquantaine d'années. Auparavant, on n'avait accès qu'à l'étude de la portion stable et reproductible des sons.

Et là, vous avez maintenant un diagramme représentant une durée sur l'espace d'une feuille de papier ?

Oui. Nous pouvons même l'obtenir sur écran, au fur et à mesure que se produisent les sons. C'est donc la possibilité de faire, en temps réel, l'association entre une forme visuelle et objective, et la forme perçue auditivement. La trace dure un certain temps, on peut aussi l'arrêter et, en tout cas, elle est plus facile à décrire que ne l'est le son. Essayez de demander à cinq ou à dix personnes de parler d'un son ; c'est un exercice très délicat qui se heurte à la difficulté de trouver les mots justes pour décrire les sons – ces mots n'ont pas le même sens pour tout le monde – et qui se heurte surtout au fait que, même s'il vous est possible de réécouter le son plusieurs fois, chacun pourra, à chaque nouvelle écoute, saisir des éléments différents du son, et il sera très difficile d'être sûr que l'on parle de la même chose.

Par exemple, il y a, sur une feuille que vous me montrez, quatre représentations, quatre dessins qui se suivent et qui paraissent se ressembler : pouvez-vous les décrire et dire ce qu'ils représentent ?

Il s'agit d'une représentation fréquence/amplitude/temps d'un fragment de parole. Les spécialistes de la parole utilisent depuis longtemps ce genre de représentation : on voit un motif qui se développe dans le sens vertical depuis le grave jusqu'à l'aigu, et horizontalement dans la dimension du temps. Il est dessiné tantôt avec des petits points, tantôt avec des raies écartées, tantôt avec des raies serrées. En fait, visuellement, on retrouve à travers ces quatre représentations quelque chose de commun.

Une sorte de colonne en particulier.

La colonne c'est la consonne *j* qui est dans le mot *aujourd'hui*. Ce mot a été prononcé par des voix différentes. Mais, à partir du moment où il a été convenablement prononcé, la forme correspondant à ce que nous sommes convenus de coder sous le mot *aujourd'hui* est commune aux quatre dessins. Elle est tout à fait identifiable comme étant le même genre de forme, alors que le substrat sonore, le matériau acoustique avec lequel elle a été produite, est tantôt du bruit, tantôt des harmoniques de fréquences différentes, tantôt un son éraillé.

Pouvez-vous me lire ces quatre *aujourd'hui*, correspondant aux quatre dessins ?

Le premier est en voix très grave appelée voix de contre-basse, le deuxième est une voix chuchotée, les troisième et quatrième sont dits par la même personne, successivement dans l'aigu puis dans le grave de sa voix. Le dessin correspondant au mot *aujourd'hui*, très reconnaissable au travers de la diversité des voix, est, en fait, la traduction des variations spectrales liées au mouvement de la langue et de tous les organes de la voix. Ce qu'on appelle des changements de formants, les interruptions du son, les bruits de consonnes occlusives et chuintantes, tous ces éléments sont extrêmement complexes si vous voulez les décrire en termes acoustiques. Or un petit enfant de trois ans comprend très bien la parole. Il est donc capable de reconnaître la forme, malgré la diversité des réalisations. Selon les locuteurs, on constate des anamorphoses spectrales (enfant, adulte, étranger) ou temporelles (vitesse d'élocution), et, si l'on n'exécède pas certaines limites, si l'on ne déforme pas trop la forme, nous sommes capables de l'extraire de ce signal complexe. De la même façon, nous procédons à de telles analyses pour en arriver à dire : « C'est une note de violon,

c'est une note de flûte, c'est une porte qui grince, c'est un bruit de clé. »

Cette capacité que nous avons d'identifier rapidement et sans erreur les sons de notre environnement quotidien semble naturelle et banale. Pourtant, il s'agit d'une conduite fondamentale dont dépend notre survie, mais qui est très mal connue. L'identification des bruits n'a pas été vraiment étudiée dans le règne du vivant. Il serait intéressant de savoir de quelle manière les animaux décodent les formes acoustiques. Tout le monde sait qu'un chat accourt lorsqu'il entend le bruit de la porte du réfrigérateur, mais est-ce seulement le contexte très limité de la situation qui lui permet de reconnaître ce bruit ? Serait-il capable de généraliser la reconnaissance du bruit de porte à tous les bruits de réfrigérateur ? Je ne sais pas... Peut-être alors inventerait-il le langage pour désigner ce bruit... Là, je m'aventure sur des terrains qui ne sont pas de ma compétence mais que je trouve passionnants dès que l'on essaie de développer cette idée de forme dans la perception sonore.

Au fond, l'approche dont vous parlez est plus qualitative que quantitative. Elle ne se contente pas de mesurer des fréquences.

Je veux bien si vous ne mettez pas une connotation dévalorisante à « qualitatif ». Dans l'esprit de certains physiciens, le « qualitatif » est moins noble que le « quantitatif ». Pourtant, on s'aperçoit aujourd'hui que l'accumulation des mesures, si elle est quelquefois utile, rend souvent difficile la tâche d'en extraire de l'information. Tous les systèmes vivants, nous le savons, sont admirablement adaptés à l'extraction de l'information. Or, détecter une structure, reconnaître une forme permet de manière très efficace de

prévoir les éléments à venir, dès lors qu'on a su mettre en mémoire des structures prévisibles, usuelles. Lorsque nous écoutons un discours, des bruits, de la musique, nous n'écoutons jamais de façon exhaustive la totalité du signal, milliseconde par milliseconde ; nous ne faisons pas l'analyse détaillée de tout le spectre ; nous sommes continuellement en train de parier sur ce qui va venir et nous faisons de petites vérifications au fur et à mesure que l'événement sonore se déroule. Si, dans certaines situations, nous sommes obligés de tout analyser, nous sommes très vite fatigués – c'est le cas à l'écoute d'une langue étrangère ou d'une musique extrêmement nouvelle et complexe –, nous sommes alors rapidement submergés par la quantité de détails qu'il faut analyser, et la fatigue vient vite. En revanche, dès que l'on se trouve dans un contexte sonore familier, avec la possibilité de détecter des structures – qui dit structure dit reconnaissance de formes –, on fait une grosse économie et on est capable de traiter une grande quantité d'informations plus vite et mieux. Cela s'apprend, bien sûr : les enfants doivent certainement passer beaucoup de temps à repérer les structures, que ce soit dans la parole, dans les bruits, dans la musique, mais, au fur et à mesure de l'apprentissage et de la mise en mémoire, ils deviennent de plus en plus compétents.

Dans votre travail, comment passe-t-on du simple au complexe ? Du son élémentaire, de la simple vibration, à un paysage sonore, une symphonie ou un objet musical très élaboré ?

C'est une question difficile. Mais s'agit-il bien d'un passage du simple au complexe ? Ne s'agirait-il pas plutôt d'un changement de niveau d'observation, d'un déplacement dans la hiérarchie du niveau d'écoute selon des échelles

temporelles différentes ? J'ai tendance à penser qu'il y a autant de complexité dans un son élémentaire d'instrument que dans la structure d'une mélodie, et que seule change l'échelle d'observation.

Prenons, par exemple, une mélodie jouée au violon. Si c'est une mélodie que je connais, je l'écoute avec un certain plaisir, mais d'assez loin, en pensant à autre chose. Mais je peux aussi, à un moment donné, m'intéresser à la manière dont la personne joue, ou aux qualités du son du violon. J'aurai alors une écoute plus proche du matériau sonore, donc une écoute à un niveau temporel plus fin, plus près des signaux élémentaires. Je vais m'intéresser à la manière dont les notes sont produites avec l'archet, à leur enchaînement, au déroulement du vibrato... Je peux encore descendre à un niveau plus fin si j'en ai la capacité. A chaque niveau d'écoute – son du violon, note musicale, mélodie –, nous structurons le signal selon des relations et dans des échelles temporelles différentes. Une bonne analogie serait celle des tableaux d'Arcimboldo qui sont des portraits d'hommes de la Renaissance composés avec des légumes et des fruits. De loin, au niveau le plus global, on voit très bien le profil d'un personnage et toutes les nuances du modelé. Lorsqu'on s'approche, pour regarder dans le détail, on découvre une poire, une tomate. Mais, à ce moment-là, on ne voit plus la forme complète. De la même façon, lorsqu'on procède à une écoute analytique du son, on éprouve beaucoup de difficultés à écouter l'interprétation musicale à un niveau supérieur. Tous ces niveaux existent, ils dépendent de la formation professionnelle, ou tout simplement de notre motivation au moment de l'écoute.

En fait, si je prends une partition musicale classique, avec des notes sur des portées, cette partition pro-

pose une forme à lire, mais elle n'est pas représentative de la forme visuelle de l'objet sonore qu'elle est censée représenter.

En effet, elle ne représente certainement pas l'objet sonore en tant que tel. Sur une partition musicale, les notes sont des signes conventionnels représentant les hauteurs discrètes de l'échelle musicale, mais pas la fréquence exacte – on peut jouer un peu plus haut ou un peu plus bas que ce qui est écrit –, et il en va de même pour les durées. La partition est faite pour être jouée, c'est un code d'action à l'intention de l'interprète. Le son n'y est pas représenté. Il peut être figuré par le nom d'un instrument – flûte ou hautbois – pour autant qu'on en ait une représentation mentale. Les musiques dont la complexité de structure réside dans des rapports de fréquences et de durées sont, finalement, assez bien représentées sous leur forme graphique : je pense à la musique de Bach qui pourrait bien en être le summum – où toute la complexité, si prenante, si captivante pour tant d'auditeurs, est contenue dans la partition, au point que certains musiciens trouvent leur plaisir à la seule lecture mentale. Maintenant, si la composition et la forme musicale se déroulent dans le domaine spectral, la partition traditionnelle est tellement schématique qu'elle devient inopérante. C'est d'ailleurs le problème de bien des musiciens aujourd'hui, qui composent avec des bruits ou des sons synthétiques, et pour lesquels la forme sonore demanderait une représentation totale, c'est-à-dire une partition fréquence/intensité/durée. Certes, les appareils acoustiques fournissent cette partition, intégrale mais... illisible. Ladite partition trop complète nous submerge sous une quantité de choses qu'il faudrait pouvoir simplifier, schématiser comme nous le faisons, apparemment, à l'écoute. Là encore, nous manque la correspondance entre la réalité

sonore et une partition qui – tout en représentant la forme totale de la musique – la traduirait en termes élémentaires, efficaces et pertinents pour l'auditeur.

Et reproductible.

Par qui ? L'homme ou la machine ?

Les deux.

Si vous voulez qu'elle soit totalement reproductible, il faut toute l'information. Or, il est certain que, lorsque nous-mêmes détectons des structures ou des formes, nous allons chercher certains éléments et en laissons tomber d'autres. Nous faisons toujours une simplification du traitement de l'information sonore. Et l'image que je vous montre là, qui est une représentation spectrale de toute une séquence composée à l'ordinateur, est beaucoup trop complexe. Ainsi, un son avec 20 harmoniques va donner lieu à 20 tracés évoluant dans le temps, alors que, pour nous, un son avec 20 harmoniques donne lieu à une sensation : on entend un son avec une certaine qualité. Nous n'avons pas, pour l'instant, les outils qui permettraient de représenter ce son à la fois de façon suffisamment complète et, en même temps, de façon simplifiée pour sa reconnaissance sur le plan de notre perception.

*

BIBLIOGRAPHIE

Castellengo M., « La perception auditive des sons musicaux », in *Psychologie de la musique* (sous la direction de A. Zenatti), Paris, PUF, 1994.

Edelman G., *Biologie de la conscience*, Paris, Odile Jacob, 1992.

Leipp E., *Acoustique et Musique*, Paris, Masson, 1989.

Pierce J. R., *Le Son musical. Musique, acoustique et informatique*, Paris, Belin, coll. « L'Univers des sciences », 1984.

Risset J.-C., « Son musical et perception auditive », *Pour la science*, 109, novembre 1986.

–, « Perception, environnement, musiques », in *Musique et Perception, Harmoniques*, 3, 1988.

Santon F., « L'acoustique des salles de concert », *La Recherche*, 211, juin 1989.

Scotto Di Carlo N., « La voix chantée », *La Recherche*, 235, septembre 1991.