



Geste – Modèle physique – Composition musicale

Claude Cadoz

ACROE – ICA
INPG, UJF, Ministère de la Culture
46, Av. Félix Viallet
38000 Grenoble
Tél : 04 76 57 46 61
Fax : 04 76 57 48 89
Mél : Claude.Cadoz@imag.fr

Le geste et la parole, plus généralement les comportements corporels et les émissions vocales sont les deux modalités par lesquelles, au-delà de ses nécessités purement vitales, l'homme est en relation avec son environnement. Le geste et la parole ont comme point commun d'être tous deux des moyens d'expression, mais à la différence de la parole, le geste permet aussi d'agir directement sur le monde physique et de le transformer matériellement. C'est probablement cette ambivalence qui fait en premier lieu du geste un sujet aussi riche et complexe. Mais une autre raison fondamentale le met aujourd'hui au centre des débats : c'est la forte perturbation que son statut a subi à l'arrivée des technologies de l'information et de la communication et le besoin d'en restaurer, dans leur contexte, toutes les dimensions. En effet, tandis que la voix, le son, l'image ont été rapidement convertibles en *signaux* et par la même sont devenus propres à la capture, l'enregistrement, la transmission, l'analyse et le traitement, le geste a vu son champ d'exercice se restreindre comme une peau de chagrin dans les relations avec ou à l'aide des nouvelles technologies. Celles-ci n'en permettaient en effet, jusqu'à récemment, qu'une prise en compte très minimale à travers des dispositifs rudimentaires comme les touches des claviers numériques, réduisant les actes à des sélections et des actions en tout ou rien.

Les questions relatives au geste sont alors bien de deux ordres : celles qui se rapportent à la compréhension de sa nature et de sa diversité, et celles qui se rapportent aux technologies et aux conditions qui permettent de les servir pleinement l'une et l'autre.

En nous plaçant dans la perspective de la création musicale qui recourt, totalement ou partiellement, aux nouvelles technologies, nous proposons ici un parcours schématique pour montrer un lien du geste naturel à la création musicale à l'aide des nouvelles technologies. Ce parcours emprunte un chemin spécifique, déterminé par l'approche même de la création artistique à laquelle nous nous sommes attachés dans nos travaux : celle de l'utilisation de l'ordinateur comme moyen de simulation de l'univers physique. En fait, nous abordons la simulation non seulement comme un paradigme pour la création de situations interactives et la production de phénomènes sensibles, mais également comme un mode d'éclairage et de compréhension général et unificateur.

Les questions que nous allons soulever adressent successivement les natures du geste, les moyens de leur prise en considération dans les relations avec ou à travers les nouvelles technologies, leurs relations avec les processus numériques de création des phénomènes sonores et des structures musicales. Nous tenterons enfin de mettre l'interaction gestuelle en perspective dans le cadre d'une utilisation générale de l'informatique pour la création musicale.



Les natures du geste

Nous reprenons ici en résumé une série de dichotomies simples mais fondamentales qui nous ont permis par ailleurs¹ de proposer un cadre typologique à l'approche du geste.

Le geste en général

Il est tout d'abord apparu nécessaire de distinguer clairement, dans les diverses invocations du geste, celles qui s'y réfèrent de manière métaphorique de celles qui s'attachent à sa réalité matérielle. Certaines dynamiques au sein d'un énoncé musical peuvent présenter des formes temporelles, des allures qui renvoient par évocation au comportement gestuel. Il faut bien noter que cette évocation peut se manifester en dehors de toute réalité effective d'un geste de production ; il suffit pour cela que des variations perceptibles de la séquence sonore présentent des similitudes d'évolutions temporelles avec celles d'un geste possible. Cela indique que notre système cognitif dispose de référents relatifs à ces dynamiques et que ces derniers peuvent jouer un certain rôle dans la structuration de nos représentations mentales. Même si le qualificatif de "gestuel" peut apporter une forme de lecture pertinente de ces dynamiques, il est clair que c'est un usage métaphorique et que l'on ne peut l'aborder avec la même méthodologie que le geste réel.

Une seconde distinction, importante dans l'observation de l'exercice du geste en amont des nouvelles technologies, s'appuie sur le geste en soi, c'est-à-dire où seul le corps est matériellement en cause, par opposition au geste qui s'effectue dans le cadre d'une interaction avec un objet physique. Le premier, dans la mesure où ses manifestations à l'égard de l'environnement ne peuvent être reçues que visuellement, ne peut être que de nature *sémiotique* (selon la terminologie que nous avons proposée [ibid.]), c'est-à-dire qu'il ne peut servir qu'à faire connaître, émettre de l'information. En musique, le geste du chef d'orchestre entre typiquement dans cette catégorie². Les comportements gestuels relatifs à la locomotion doivent être exclus de cette catégorie dans la mesure où le corps interagit avec un objet physique (certes particulier) : le sol avec ses reliefs. Dès lors qu'un objet physique intervient, deux circonstances nouvelles apparaissent : (1) le système physique en jeu est l'ensemble constitué du corps et de l'objet ; ses propriétés et donc ses états et ses évolutions sont différents de celles du corps seul, (2) selon la nature de l'objet, le comportement gestuel peut servir différents objectifs. Il est alors pratique de distinguer deux catégories très différentes selon que l'objet en question interagit matériellement avec l'environnement physique et permet alors de le modifier ou de le transformer (comme une pelle, une pioche ou un marteau) ou qu'il est spécialement conditionné pour produire ou relayer des phénomènes sensibles. On pourra parler d'outils dans le premier cas, d'instruments dans le second. La catégorie qui

¹ CADOZ C., "Continuum énergétique du geste au son : simulation multisensorielle d'objets physiques". in "les interfaces pour la création musicale". – Dir. H. Vinet, F. Delalande. HERMES Editeur. 1999.

CADOZ C., "Musique, gestes, technologies". in "Les nouveaux gestes de la musique", sous la direction d'Hugues Genevois et Raphaël de Vivo. 1999.

CADOZ C., WANDERLEY M., "Gesture and Music". in Trends in Gestural Control of Music. IRCAM Editeur. 2000. avec CDROM.

² À noter que lorsque le chef d'orchestre a une baguette, il y a bien une interaction avec un objet physique. Le statut de cette situation pourra être clairement défini avec des critères plus fins introduits plus loin.



retiendra notre attention ici est naturellement la seconde mais on peut signaler au passage que ces deux catégories ont en commun le fait qu'une relation énergétique (*ergotique*) s'établit entre le corps et l'objet : indépendamment de la finalité matérielle ou communicationnelle, une dépense effective d'énergie en dehors du corps a lieu, qui met en mouvement ou déforme l'objet physique. Ce dernier peut alors transmettre cette énergie à un tiers (par exemple l'environnement aérien dans lequel se propagera une onde acoustique) ou la restituer partiellement au manipulateur qui, de ce fait dans le même temps reçoit une forme d'information à travers ses capteurs tactiles et proprio-kinesthésiques. Le canal gestuel est dans ce cas un moyen de connaissance (fonction *épistémique*) qui s'ajoute aux autres sens. Notons que le geste instrumental a ceci de particulier qu'il met en jeu de manière indissociable les trois fonctions, *sémiotique*, *ergotique* et *épistémique* que nous avons introduites au fil de ce paragraphe. C'est là sans doute une de ses propriétés les plus importantes. Deux remarques s'imposent alors à ce niveau :

- Le lien entre la fonction *sémiotique* et la fonction *ergotique* peut être très fort : certaines figures dynamiques expressives ne peuvent s'obtenir qu'à la condition que le geste soit pleinement physique (donc *ergotique*). L'aspect *ergotique* est donc partie prenante à part entière du processus communicationnel et il est illusoire de penser que toutes les formes d'information gestuelle peuvent être obtenues à main nue puisque la physique du système dans le cas de la situation instrumentale fait bien intervenir au moins deux composants : le corps humain et l'instrument, et les propriétés de l'ensemble ne peuvent s'obtenir et se décrire qu'en considérant l'ensemble.
- Le lien entre la fonction *sémiotique* et la fonction *épistémique* est également très fort puisque, là aussi, il s'agit d'un système, d'une inter-relation : la conduite des mouvements expressifs se fait, pour certains gestes, dans la logique d'une boucle intime où chaque phénomène de perception tactile et kinesthésique pendant l'action est susceptible de modifier dynamiquement la suite du geste.

Dans la catégorie des instruments, une dernière dichotomie est utile : celle qui permet de mettre d'un côté les instruments qui, du geste humain à la perception, s'appuient sur une continuité énergétique et ne font appel (ce qui est un corollaire) à aucune source d'énergie extérieure à l'homme, et de l'autre les dispositifs qui interviennent pour moduler un flux énergétique (par exemple lumineux) qui s'établit entre une source extérieure et nos sens (par exemple la vision). Le geste instrumental musical traditionnel (c'est-à-dire avant l'électronique et mis à part quelques cas singuliers comme l'orgue) se place dans le premier cas. Le geste du marionnettiste entre dans le second cas puisque les mouvements de la marionnette ne sont visibles que du fait de son interaction avec le faisceau d'un projecteur. Le geste d'écriture (musicale comme littéraire) ou le geste graphique entrent également dans cette seconde catégorie : le crayon ou le pinceau sont des instruments ;, mais le dispositif se complète par une fonction nouvelle que l'on peut appeler d'une manière générale la *gravure* qui est essentiellement la transformation d'une forme temporelle en une forme spatiale. Cette fonction essentielle est la base de tous les processus de mémorisation des phénomènes temporels. Elle se généralise lorsque la trace n'est plus nécessairement directement visible ou lisible sans un intermédiaire, comme dans la gravure magnétique ou "micro-optique" des CD.

En ce qui concerne la musique, ce premier tronçon du parcours nous permet de distinguer au moins quatre types de situations, et donc quatre fonctionnalités différentes du geste : le geste au sens de la métaphore, le geste au sens du chef d'orchestre, le geste dans l'écriture et le geste instrumental. Avant d'aborder le problème du geste dans le cadre des nouvelles technologies, il faut accorder quelque attention particulière au geste instrumental musical.



Le geste instrumental musical

La fonction physique que doit remplir l'instrument de musique est précise : elle est quantitative et qualitative. Qualitative, elle est de permettre la transformation de phénomènes gestuels en phénomènes acoustiques. Quantitative, elle est de permettre, à travers cette transformation, l'émission d'une certaine énergie à destination des auditeurs à partir d'une énergie développée par l'instrumentiste. La seconde fonction ne peut être remplie qu'à la condition que la relation entre l'instrumentiste (qui est la seule source d'énergie) et l'instrument soit une interaction physique. Dès lors, toute action de l'instrumentiste sur l'instrument est indissociable d'une réaction du second sur le premier et l'ensemble instrumentiste-instrument forme en soi un système. Les phénomènes d'interaction dépendent des propriétés de l'un et de l'autre et c'est par la même, en particulier, que l'instrumentiste, à travers ces différents phénomènes, reçoit par ses sens tactiles et proprio-kinesthésiques une certaine information sur les propriétés et les comportements intimes de l'instrument.

Selon le point de vue quantitatif, l'instrument est donc une chaîne énergétique et la première forme du geste instrumental est déterminée par l'objectif de fournir efficacement à l'instrument une énergie qu'il pourra convertir en énergie acoustique. Nous avons nommé cette contribution gestuelle "*geste d'excitation*" [ibid.].

La chaîne matérielle (chaîne instrumentale élémentaire) assurant la transmission énergétique fait appel à des composants matériels qui présentent, par nécessité et par utilité une certaine permanence. En conséquence de quoi, les solutions adoptées pour assurer cette transmission détermineront autant de catégories fondamentales d'instruments. Mais pour un même principe physique de transmission, de multiples variations sont possibles et certaines d'entre elles sont même envisageables à partir d'un même dispositif par modulation ou modification de quelques propriétés simples de certains composants (longueur d'une corde ou d'une colonne d'air par exemple). Les actions gestuelles peuvent être impliquées dans cette tâche de modification qui aura pour effet de moduler le phénomène de génération sonore pendant le jeu instrumental même. Cette seconde composante du comportement gestuel, qui ne s'inscrit pas de la même façon dans le processus énergétique mérite d'être identifiée pour elle-même. Nous l'avons nommée "*geste de modification*".

Enfin, plusieurs dispositifs différents ou correspondant à des variantes physiques d'un même principe peuvent être impliqués simultanément, dans un instrument complexe) pour remplir en séquence ou en collaboration la fonction principale de production sonore. L'instrumentiste s'implique alors gestuellement d'une troisième façon, dont la physique est distincte des deux précédentes. Nous appelons "*geste de sélection*" ce qui correspond à ce choix en ligne, pendant le jeu même, de la ou des chaînes instrumentales. Pendant la sélection (comme dans le cas du jeu pianistique par exemple) l'interaction matérielle avec l'objet peut s'interrompre. Dans le cas nominal, le geste de sélection est donc à main nue et n'entre pas dans le cadre *ergotique*.

Ce rapide parcours montre que la catégorie même du geste instrumental peut être en soi multiforme puisque, construite autour de cette condition énergétique centrale, elle met en jeu les fonctions gestuelles de manières différentes et combinées : la fonction *ergotique* est déterminante dans la composante "*excitation*", elle y est évidemment indissociable des fonctions *épistémique* et *sémiotique*. Dans la composante "*modification*", la fonction *ergotique* peut être présente (comme lorsqu'il s'agit de tendre la peau d'une timbale à l'aide d'une pédale) ou quasiment pas (comme lorsqu'il s'agit de modifier la longueur d'une corde



dans le violon). Pour la composante “*sélection*”, il n'existe pas, hormis dans les automates musicaux, de dispositifs chargés de gérer ou d'assister cette sélection. La composante “*sélection*” se caractérise alors par le fait que ce sont pratiquement exclusivement les contraintes physiques et physiologiques humaines qui président à leur conception. En revanche, la fonction *épistémique* du canal gestuel peut y être mise fortement à contribution, lorsqu'il s'agit par exemple de recalibrer à chaque instant le positionnement de la main sur un clavier en s'aidant du contact avec les touches enfoncées ou les touches noires voisines.

Ces quelques considérations de base permettent d'amorcer une méthodologie générale pour parcourir l'espace gestuel et identifier ses différentes natures. On peut constater au passage que le développement de cette analyse se fait en parallèle avec une incursion de plus en plus intime dans la structure même de l'instrument. C'est dire qu'en fait, l'analyse du geste instrumental et l'analyse de l'objet instrument sont indissociables et qu'il serait vain, puisqu'il s'agit fondamentalement d'une interaction, d'envisager l'une sans l'autre.

Un parcours complet de toutes les lignées, espèces et individus instrumentaux qui ont résisté à la sélection humaine serait fort long. Aussi, on se contentera ici d'un seul niveau complémentaire dans l'arborescence : celui des formes d'excitation, pour montrer au passage que si les catégorisations procèdent indubitablement de choix culturels complexes, elles résultent également de contraintes propres aux règles de l'univers dans lequel nous vivons.

C'est le cas en effet pour la catégorisation des instruments et des gestes d'excitation correspondant en ces deux grandes classes que sont les instruments à excitation instantanée (percussion, cordes frappées, pincées, etc.) et les instruments à entretien continu (à frottement d'archer, à vent, etc.). Il n'existe en effet, “dans la nature”, que deux façons de communiquer de l'énergie à une structure vibrante à partir d'une structure non vibrante (comme le corps humain), celle où le transfert se fait quasiment instantanément, séparant le geste et le son par une frontière événementielle ponctuelle (fin de l'un, début de l'autre) et celle où le transfert se fait en continu, moyennant une disposition physique très particulière (comme on peut l'analyser entre le crin collophané et la corde) et où geste et son se développent dans le même temps. Les gestes d'excitation correspondant sont évidemment de deux natures complètement différentes, difficiles à ramener l'une à l'autre. Ce qui permet de dénoncer au passage l'ineptie des instruments électroniques qui offrent à jouer des sons de violon en enfonçant les touches d'un clavier.

La création sonore à l'aide des nouvelles technologies

Dans le cadre de la création sonore à l'aide des nouvelles technologies, le geste instrumental ne répond plus à aucune nécessité matérielle. En effet, le propre des systèmes électroniques quels qu'ils soient est d'introduire, par le principe de l'amplification, une coupure fonctionnelle fondamentale dans les circuits énergétiques. Le transistor (et avant lui la triode sont des relais : ils mettent en jeu deux circuits énergétiques indépendants, dont l'un (le circuit de commande) à faible énergie, *module* les conditions de l'autre (le circuit actif) à énergie élevée, à la manière d'un robinet dont on ajuste l'ouverture pour contrôler le débit d'eau. Les quantités d'énergie dans l'un et l'autre circuit peuvent être d'échelles très différentes, en particulier l'énergie dans le circuit de commande peut être négligeable devant celle mise en jeu dans le circuit actif. Par ailleurs, sauf à le provoquer expressément par un dispositif de “*feedback*”, il n'y a pas d'échange d'énergie entre les deux circuits. Enfin, si le circuit de commande peut déterminer le flux énergétique du circuit actif, l'inverse n'est pas vrai : la relation est unilatérale.



Le principe du relais, qui peut prendre des formes matérielles très variées est, par son exploitation systématique et généralisée dans les “nouvelles” technologies, l’un des quelques principes technologiques de base qui permettent de les caractériser par rapport aux “anciennes”, même s’il est déjà présent depuis l’Antiquité dans des machines comme le premier clavier d’orgue du mécanicien d’Alexandrie Xtésibios en 270 avant J.C. Les notions de signal et d’information elles-mêmes lui doivent l’existence. C’est en effet à partir du moment où autre chose que de l’énergie ou de la matière peut être véhiculé entre deux lieux que l’on parle alors d’information, portée par un signal. Ce sont ces conditions matérielles qui déterminent également les aspects relationnels : à la différence d’une grandeur énergétique ou matérielle, une entité informationnelle ne se véhicule que dans un sens. Dans la même logique, les entités informationnelles peuvent être “diffusées”, c’est-à-dire distribuées sans se diviser à plusieurs destinataires (ce qui n’est évidemment pas le cas pour la matière et pour l’énergie). Enfin, les conditions énergétiques à l’émission et à la réception peuvent être totalement indépendantes : l’énergie nécessaire à l’émission peut être faible et les énergies mises en jeu au récepteur très importantes.

Il est clair que dans ce contexte, la première transformation qu’opèrent les nouvelles technologies dans le domaine de la création sonore portent, de manière profonde, sur les conditions matérielles de la production des phénomènes sensibles à partir des actes physiques humains (gestuels). En conséquence, la chaîne de production dans le cadre des nouvelles technologies ne peut plus être analysée selon les mêmes critères et les systèmes électroniques et informatiques dédiés à la création sonore ne peuvent être définis comme procédant simplement d’une “nouvelle lutherie”.

En restant au niveau de la production du matériau, par opposition à celui de la création et de la maîtrise de la structure musicale (compositionnelle), l’évolution des approches marque schématiquement trois grandes étapes : une première approche que l’on pourrait qualifier de “phénoménologique”, une seconde “fonctionnelle” et une troisième “structurelle”.

La première, dont le représentant historique est la démarche de la musique concrète, se caractérise en fait par l’appropriation du signal (sonore) en tant qu’objet et par le développement de toutes les techniques de son traitement et de ses transformations (segmentation, transformations temporelles, fréquentielles, morphologiques, chaînage, superposition). Ces techniques sont le substrat matériel de la construction de “langages”, ou au moins de processus de création musicaux spécifiques. Elles se prolongent dans le cadre des technologies numériques et trouvent une forme synthétique et générale dans toutes les approches fondées sur le traitement du signal.

La seconde, qui a vu le jour en particulier avec la synthèse sonore par blocs fonctionnels représentée par le “paradigme Music V” de Max Mathews va au-delà de la surface (phénoménologique) du signal. Elle met en œuvre, en amont, des processus fonctionnels dont l’agencement en structures hiérarchisées (fondées sur la notion de commande de blocs fonctionnels par la sortie d’autres blocs fonctionnels de même nature) engendre le signal avec toutes ses propriétés. Cette approche peut d’ailleurs s’arrimer à la précédente dans la mesure où les signaux engendrés peuvent être également traités. Le substrat des langages et de l’expressivité sonores est dans ce cas déterminé par la base fonctionnelle des modules du système. Dans le cas de Music V est des systèmes qui en dérivent et abritent les différentes méthodes de synthèse développées depuis les années 60 et 70 (synthèse additive, soustractive, FM, FOF, granulaire, etc.), appelés généralement aujourd’hui systèmes de synthèse en modèle



de signal, ces bases relèvent toutes d'un même angle de vue : le phénomène sonore est une entité en soi, décomposable en constituants élémentaires caractérisés par des paramètres (qu'il s'agisse des sinusoïdes ou des grains de Gabor) et tout le problème de la génération sonore est ramené aux différentes manières d'engendrer ces éléments ou des ensembles structurés de ces éléments, et de contrôler leurs paramètres ou des ensembles structurés de leurs paramètres.

La troisième approche, historiquement la dernière, pose le problème d'une façon différente. Représentée par les courants relevant du "modèle physique", elle déplace l'objet d'étude en amont du phénomène, au niveau de sa cause. Elle s'approprie l'objet plutôt que le signal. Alors que les deux approches précédentes partent du signal respectivement pour le traiter ou pour l'engendrer, celle-ci part de l'objet et le simule. Il serait exagéré de caractériser aussi clairement toutes les approches qui se réclament du modèle physique car nombre d'entre elles n'invoquent que très marginalement le paradigme de la simulation et restent très largement dans un environnement général d'exploitation basiquement conditionné par les méthodes en modèle de signal. Mais toutes les approches "modèle physique" en revanche ont comme trait commun de se fonder, en allant plus ou moins en profondeur, sur une analyse structurelle des causes physiques sonores naturelles et une construction des simulacres selon une structure qui tend à suivre à l'identique la structure de l'objet de référence. Ainsi, par exemple, pour réaliser la synthèse d'un son de guitare, on mettra en jeu dans la simulation des composants algorithmiques correspondant respectivement aux cordes, au chevalet, à la caisse, etc. Dans ce troisième cas, la référence aux objets et phénomènes réels intervient comme élément structurant du vocabulaire sonore.

L'approche modèle physique, dans la mesure où elle conduit de toute manière à la génération d'un son numérique se raccorde évidemment aux méthodes de synthèse du son et de traitement du signal.

L'histoire des technologies récentes pour la musique présente, singulièrement, un cheminement inverse de celui de l'évolution des instruments eux-mêmes : on part en effet de la destination (le son) pour remonter à son origine (l'instrument). Ceci démontre d'une autre manière que les nouvelles technologies ne peuvent être l'occasion de nouvelles lutheries, qu'elles se placent dans une position non pas de prolongement ou d'extrapolation mais de "réflexion" (au sens du miroir). Bien sûr, le miroir est actif et c'est là sans doute ce que l'ordinateur instaure, sans équivalence antérieure : la possibilité de créer des "images" des représentations reflétant pour parties les propriétés du monde réel, mais capables aussi d'y intégrer des propriétés nouvelles. Resterait ensuite à voir comment ces images peuvent éventuellement susciter la création de leurs originaux réels.

Le geste et la machine

La "machine" d'aujourd'hui est informatique ou en tout cas au moins "informatisée", ce qui signifie que les phénomènes physiques dont elle est le siège sont, indépendamment des fonctions qu'ils remplissent, de nature électronique et numérique. L'homme, quant à lui, est toujours biologique, c'est-à-dire que les phénomènes physiques par lesquels il est en relation avec ce qui l'entoure (par exemple une machine) sont des ondes acoustiques, des ondes lumineuses, ... pour ses récepteurs et des mouvements mécaniques pour ses comportements et perceptions gestuels. Interagir gestuellement avec une machine suppose donc expressément une fonction qui n'avait pas lieu d'être dans les technologies antérieures, puisqu'il s'agit de transformer des phénomènes mécaniques (ou pneumatiques lorsqu'il s'agit du souffle) en phénomènes électroniques numériques. Cette fonction de "transduction" (littéralement transformer / conduire) s'est posée pour tous les domaines d'application de l'électronique. Par



exemple, il a fallu créer des transducteurs acoustico-électriques pour transformer les vibrations acoustiques en signaux électriques ou des transducteurs électro-acoustiques (les... hauts parleurs) pour la transformation inverse. La première condition pour exercer un geste à l'adresse d'une machine est donc que celle-ci dispose de transducteurs gestuels.

Les technologies des transducteurs gestuels sont variées, mais néanmoins déterminées par un petit nombre de principes. Un premier critère va les séparer en deux catégories principales selon la nature du phénomène pris en considération pour représenter le geste. En effet, le geste est un phénomène par nature mécanique, mais il peut aussi se manifester, au moins partiellement, sous forme visible. En conséquence de quoi l'on peut décider de capter des mouvements mécaniques ou des variations lumineuses (c'est ce que fait l'homme dans le sens de la "perception gestuelle"). Dans le premier cas, un intermédiaire matériel est nécessaire, avec lequel doit s'établir un contact physique direct ; la relation gestuelle pourra être, au moins localement, instrumentale. Dans le second cas, aucune interaction physique ne s'impose. Il s'appuie typiquement sur la capture visuelle, par exemple par camera, et sur une ensemble de technologies d'extraction de données comme les contours de la main, des membres, du corps de l'opérateur et, à partir de celles-ci, de variables représentatives du comportement corporel³. Loin de correspondre à une catégorie secondaire, il ne permet cependant d'exploiter que la fonction *sémiotique* du geste. Le geste concerné est en conséquence exclusivement non-instrumental.

Les technologies de la transduction mécanico-électrique sont, quant à elles multiples. Elles sont toutes basées sur la production ou la modification d'un phénomène électrique en conséquence au déplacement ou à la déformation d'un objet (que celui-ci soit solide ou fluide). Le cœur de la fonction de transduction repose sur quelques principes, en nombre limités : variation de conduction électrique (d'effet résistif), variation d'effet inductif (déplacement d'un corps magnétique à proximité ou à l'intérieur d'une bobine ou d'antennes), variation d'effet capacitif, effet piezo-électrique, effet de champ, pour les principaux. On peut y ajouter les variations de transmission d'ondes ultrasonores, lumineuses, ou électromagnétiques en général.

Indépendamment de la technologie de transduction proprement dite, un autre critère intervient à ce niveau, qui partage à son tour les dispositifs en deux catégories selon que l'objet soumis à déplacement ou déformation est attaché au corps et se présente en quelque sorte comme une prothèse ou un prolongement de celui-ci, ou qu'il est face à l'opérateur qui peut choisir d'entrer en contact avec lui ou non. Dans le premier cas, on parle de techniques immersives, dans le second de techniques de vis-à-vis. La relation avec l'objet et ses significations possibles ne sont pas les mêmes. Dans l'une, la technologie prolonge l'homme, dans l'autre elle prolonge son environnement.

³ Un dispositif célèbre auquel il faut rendre hommage car il fut le tout premier, en 1920, et celui qu'inventa Theremin. Exploitant un effet capacitif entre la main et deux antennes sphériques, il permettait le contrôle sans contact de la fréquence et de l'amplitude d'oscillateurs électroniques. Contrôlant des grandeurs relatives à la « substance » sonore, il n'est cependant pas instrumental, du fait de l'absence de contact. Il est le premier exemple historique introduisant une rupture entre la nature du geste et la nature des phénomènes contrôlés.



Dans la première catégorie, les divers dispositifs développés et exploités aujourd'hui se rangent à leur tour dans quelques catégories de base⁴ représentées par le gant et les combinaisons de données (*dataglove*, *datasuit*), les exosquelettes de la main, du bras, du corps, etc. Il faut mettre dans ces catégories le cas des objets "embarqués" par la (ou les) main(s) ou le corps et qui portent soit des sources (lumineuses, infra lumineuses, infra ou ultrasonores, électromagnétiques) émettant vers des capteurs situés dans l'enceinte du jeu corporel, soit, à l'inverse, des récepteurs de ces mêmes catégories de phénomènes. Dans tous les cas, les informations captées et transmises sont des représentations des comportements corporels intrinsèques et non d'une interaction physique réelle, même si parfois la lourdeur des systèmes "embarqués" peut modifier la dynamique et les trajectoires des mouvements. Il s'agit donc, malgré la présence et le contact avec un objet physique, d'une relation purement *sémiotique* et donc non-instrumentale.

Dans la seconde catégorie, la plus simple et la plus ancienne est l'interrupteur, utilisé dans le manipulateur de télégraphe aussi bien que dans les claviers alphanumériques d'aujourd'hui. L'image du geste, en tout cas dans sa fonction *d'excitation* telle que nous l'avons définie, est réduite à sa plus simple expression : instant de contact et quantification en tout ou rien. Mais une série d'interrupteurs est en revanche bien adaptée à la capture de la composante de *sélection*. C'est cette dimension exclusive que les claviers d'instruments électroniques ont exploité essentiellement pendant longtemps. Le déplacement d'un curseur sur une piste conductrice, que celle-ci soit rectiligne ou circulaire, donne une image moins squelettique du geste puisqu'elle peut porter des variations "analogiques". Les évolutions gestuelles sont toutefois réduites à une seule variable. L'"habillage" physique du curseur (bouton à déplacement rectiligne ou rotatif, touche à grand enfoncement, bras articulé, etc.) permet d'exploiter ce principe dans une gamme de situations variées pouvant s'approcher de situations de référence instrumentales naturelles. Avec un habillage minimal, ce type de transducteur est un candidat raisonnable au support de la fonction de *modification* (dans sa modalité analogique, par opposition à "discrète" [ibid.]). Mis en batterie, il peut de plus jouer pour la "sélection-de-modification". Il faut en revanche que l'habillage en question soit un peu élaboré, en incluant par exemple une mécanique spécifique (dont la plus simple est le ressort de rappel), pour se rapprocher de quelques conditions naturelles élémentaires du geste *d'excitation*. Il faut bien noter que ces conditions seront alors figées pour un dispositif donné.

C'est dans cette seconde catégorie, avec une disposition nouvelle, qu'il faut situer les transducteurs gestuels rétroactifs (TGR). Ils peuvent exploiter l'une quelconque des technologies précédentes pour capter les phénomènes d'interaction opérateur-objet physique, mais ils portent une fonction symétrique supplémentaire : des transducteurs électromécaniques (en d'autres termes des moteurs électriques) qui peuvent alors assurer un retour de la machine vers les capteurs tactiles et proprio-kinesthésiques de l'opérateur. L'ensemble des caractéristiques d'un TGR est alors déterminé par celles de ses trois composants fonctionnels : les capteurs, les effecteurs et son "habillage". Les premiers et les seconds sont déterminés quantitativement par les caractéristiques de précision, dynamique et bande passante. Elles doivent être adaptées à celles des variables gestuelles en cause : les déplacements gestuels peuvent être significatifs dans une plage importante (du micromètre au mètre), les efforts peuvent aller de quelques centièmes à quelques centaines de newtons. La

⁴ On pardonnera cette énumération purement catégorielle et trop rapide pour évoquer nommément les dispositifs ingénieux et fort pertinents de nombreux collègues dont nous apprécions tout particulièrement le travail.



bande passante des phénomènes gestuels (en force et en déplacement) peut aller de 0 jusqu'à plusieurs KHz (à noter qu'il a fallu un certain nombre d'années pour prendre conscience de cette nécessité et que nombre de systèmes à retour d'effort ont pâti longtemps de la croyance que quelques dizaines de Hz suffisaient). L'"habillage" enfin est une caractéristique importante. Il s'agit schématiquement des conditions géométriques (morphologie et trajectoires) et dynamiques selon lesquelles l'objet physique réel (et incontournable) se présentera à la saisie et à la manipulation de l'opérateur. Enfin, un TGR peut être constitué d'un ensemble de chaînes capteurs-effecteurs conférant au tout un certain nombre de degrés de liberté. Rappelons que la main, en tant que système mécanique, comporte à elle seule 23 degrés de liberté. À noter également que l'habillage peut être envisagé comme s'appliquant non pas à chaque DDL terme à terme, mais à des groupements structurés de DDL, pour permettre par exemple des manipulations de type joystick, pince, déplacement d'un solide tridimensionnel, etc.

La technologie des TGR pose un problème particulier, sans équivalent pour les autres transducteurs (visuels ou acoustiques) : du fait de la nécessité d'un intermédiaire matériel réel, il ne peut exister de dispositions universelle permettant la prise en charge de toutes les catégories d'interactions physiques possibles. Les TGR sont alors nécessairement catégorisés en fonction de leur application. Néanmoins, il est possible, comme le propose la technologie que nous avons développée à l'ACROE⁵ de mettre en jeu une forme de généralité fondée sur la modularité des trois composants de base (en particulier l'habillage).

L'introduction de la rétroaction permet de prendre en charge la fonction *ergotique* du geste et, par conséquent de la composante *excitation* du geste instrumental d'une façon non plus figée, mais programmable, puisque les comportements mécaniques dynamiques de l'objet manipulé pourront être contrôlés par la machine, dans une boucle d'interaction effective. Cette boucle pourra associer, par des algorithmes appropriés, les données d'entrées gestuelles à la production de comportements mécaniques s'adressant aux capteurs tactiles et proprio kinesthésiques de l'opérateur.

Systemes gestuels et synthèse

Nous disposons maintenant, avec les processus de synthèse et les transducteurs gestuels, de tous les ingrédients permettant d'établir la relation du geste à l'oreille en utilisant un ordinateur. Il existe naturellement un grand nombre de façons d'associer les premiers et les seconds, dans la mesure où la seule condition pour qu'il se passe quelque chose est que les variables issues des systèmes gestuels et celles que les dispositifs de synthèse requièrent se correspondent, au moins quantitativement. Toutefois, notons que pour la prise en compte de la rétroaction, donc pour soutenir le geste *d'excitation* en particulier (assurer la fonction *ergotique* en général), les dispositifs de synthèse doivent également produire les signaux de "sortie gestuelle" (les signaux de commande des moteurs de rétroaction).

⁵ CADOZ (C), LUCIANI (A), FLORENS (JL), 1984, « Responsive Input Devices and Sound Synthesis by Simulation of Instrumental Mechanisms : The Cordis System », Computer Music Journal, 8, N°3, pp. 60-73. M.I.T. Press, Cambridge Mass.
CADOZ (C), LISOWSKI (L), and FLORENS (JL), (90) - "A Modular Feedback Keyboard Design", Computer Music Journal, Vol. 14, No.2, Summer 1990.



La relation entre les données gestuelles et les données de synthèse peut être posée en termes banals de “contrôle” et de “mapping”. Contrôle au sens où l’on considère que le son est le produit d’un algorithme quelconque et que les paramètres d’entrée de celui-ci sont des variables de statut non nécessairement déterminé, ou répondant aux paramètres acoustiques classiques. “Mapping” dans le sens où l’on considère les aspects relatifs au geste et relatifs au son séparément et indépendamment dans un premier temps, pour établir leurs correspondances dans un second temps. Les critères de correspondance peuvent alors évidemment être larges et ne s’appuyer que sur les caractéristiques des signaux analysés de manière neutre. Cette approche laisse le champ ouvert à toutes les utilisations possibles, mais au-delà d’un tel objectif de généralité, nous préférons développer une approche guidée par la recherche de ce qui peut rendre le geste et la production sonore cohérents entre eux.

Systèmes gestuels et simulation des objets physiques

La synthèse sonore par modèles physiques particulières permet d’aborder dans un même principe, et dès le niveau élémentaire, la création des phénomènes sonores et des phénomènes mécaniques correspondant aux comportements des objets au niveau gestuel. La particule inertielle élémentaire (définie comme une masse ponctuelle dans un espace physique) et l’interaction viscoélastique (linéaire ou non-linéaire) sont en effet deux composants de base qui permettent, par leur combinaison en réseau de forme quelconque, de simuler des mécanismes aux comportements vibratoires et d’établir la correspondance physique requise entre des variables (forces ou déplacements) représentant les grandeurs relatives aux actions gestuelles, et les variables (déplacements ou forces) représentant les grandeurs relatives aux réactions de l’objet. Des variables gestuelles aux variables sonores (et éventuellement d’ailleurs aux variables visuelles, dans le cadre d’une multisensorialité généralisée), il n’y a pas de changement de système de représentation, seulement des changements d’échelles et de propriétés, les unes et les autres pouvant être abordées à l’aide d’un même langage général et d’un même nombre restreint de concepts.

L’ensemble de la chaîne instrumentale peut alors être mis en représentation, simulé sous la forme d’un instrument virtuel. Cet instrument virtuel, mécanique dans sa conception, pourra alors être “joué” à l’aide d’un ensemble de transducteurs gestuels, parmi lesquels on disposera des TGR, de façon à réaliser une relation instrumentale. La relation gestuelle avec l’instrument virtuel pourra ainsi être réellement instrumentale.

Nos travaux concernant la musique, à l’ACROE, s’attachent à cette situation, en mettant en œuvre, à l’issue des travaux de recherche correspondant, un langage de modélisation et de simulation des objets physiques par réseaux particuliers, le langage CORDIS [ibid.], des transducteurs gestuels rétroactifs, en particulier à partir de la technologie du Clavier Rétroactif Modulaire [ibid.], un environnement de création, l’outil GENESIS.

À l’aide de CORDIS et GENESIS⁶, nous avons réalisé la modélisation de la plupart des catégories de dispositifs instrumentaux, en nous fondant sur une décomposition de la chaîne instrumentale en ses composants de base : exciteur, structure vibrante, chevalets-tables, environnement de rayonnement, etc. Pour chacun de ces composants, les catégories de base

⁶ CADOZ (C) “Musical creation with Multisensorial Interactive simulation of Physical Objects” - SCI 2001 – Orlando – à paraître – 2001.



ont été modélisées : structures vibrantes simples (cordes, colonnes d'air, plaques, membranes, volumes, ensembles de particules. etc.), modes d'excitation par percussion, pincement, entretien continu (archet, anches, etc.), caisses et tables de résonance, milieux propagatifs simples... Les résultats se concrétisent par un instrumentarium large permettant de créer la plupart des phénomènes sonores naturels et instrumentaux, ainsi que de créer des objets de nature physique mais introuvables dans la nature. Il est possible également, par cette approche, d'envisager la création sonore en partant de l'idée du son voulu ou en référence, a priori, et en déterminant, grâce à une méthodologie précise quel processus mécanique (réaliste ou non) permet de le produire. L'environnement permet naturellement de créer des ensembles instrumentaux qui, lorsque les machines sont assez puissantes, les TGR et les instrumentistes respectivement en nombres suffisants, peuvent être joués en exécution d'une partition ou en improvisation comme s'il s'agissait d'instruments réels.

En attendant que les machines atteignent (ce qui n'est pas une utopie) les puissances requises pour que le temps-réel soit applicable à une échelle significative, nous terminons ce parcours avec la mise en perspective du processus de création musicale tel qu'il peut être d'ores et déjà pratiqué aujourd'hui avec les moyens actuels.

Geste – modèle physique particulière – Composition musicale

Simuler "aussi" l'instrumentiste

L'instrumentiste est un être humain, intelligent et sensible. Mais, sans que ceci n'enlève rien à cela, c'est aussi un système physique. À ce titre, la partie qui ne concerne que ce dernier aspect peut être (moyennant quelques simplifications) modélisée et simulée, avec par exemple un langage tel que CORDIS et un environnement tel que GENESIS.

Les figures gestuelles sont d'une variété infinie : gestes des doigts qui pianotent, qui grattent ou pincet, gestes de la main qui attrape, tire, pousse, caresse, frappe, pose, gestes des bras qui bercent, moulinent, lancent, gestes du corps qui trépigne, sursaute, balance, etc. Mais quelques figures de base, à condition qu'elles soient pratiquées par les bons éléments ou les bons ensembles, permettent de les engendrer plus ou moins toutes. Ce sont les figures respectivement du *déclenchement de l'accompagnement* et de la *balistique*.

Le déclenchement résulte de la décision, à un moment donné, d'entrer en action avec une intensité variable, depuis la simple pichenette jusqu'au frappé violent. Deux grandeurs physiques peuvent en représenter les caractéristiques : la vitesse et l'inertie. Alors, une masse ponctuelle d'une inertie donnée, lancée avec une certaine vitesse initiale, peuvent jouer le rôle, au moins pendant un moment, de la partie d'un instrumentiste qui voudrait pincer, percuter ou pousser quelque chose. Il suffit alors, pour simuler l'instrumentiste à ce moment et pour cette action, de simuler une telle masse et de la faire interagir avec les objets d'une scène instrumentale qui pourraient se prêter à ce jeu.

L'accompagnement et la balistique sont deux formes semblables par un point, elles durent et la nature de leur évolution est leur substance – et différentes par un autre, chaque instant du premier est voulu et contrôlé, tous les instants de la seconde sont déterminés une fois l'action commencée. Chez un pianiste, lorsqu'il a lancé son corps, séant en bascule sur le siège, buste en balance au-dessus, épaules en roulement, bras et avant-bras en fléau dans leur prolongement pour exécuter un trait de Chopin, l'impulsion, dès qu'elle est lancée, du bas du



dos, se propage comme une onde jusqu'au crépitement des doigts sur l'ivoire. Cette colonne de propagation peut être modélisée comme un assemblage de composants inertiels en interaction viscoélastique (avec toutes nos excuses auprès des pianistes). Plus qu'une simple corde, dont il faudra toutefois que les fréquences propres soient de l'ordre des fréquences gestuelles (basses, de 0 à 20 Hz), une sorte de chaîne se dédoublant puis se décuplant, possède la balistique d'un tel geste.

On voudra moduler cette balistique et garder le contrôle des doigts, un par un, il nous faudra recourir, pour les combiner astucieusement, à nos déclencheurs précédents. On voudra, encore, différemment, frotter par un geste souple et modulé, il nous suffira de "lisser" quelques impulsions judicieusement envoyées par les mêmes déclenchements avec un peu de coefficient de frottement visqueux au bon endroit.

L'instrumentiste réel, loin d'être détrôné, trouvera là une nouvelle place. Au jeu réel sur des instruments réels, il peut ajouter le jeu réel sur des instruments virtuels, et aussi jouer d'une nouvelle façon en duo, trio, ... en "écoutant" ses partenaires (virtuels) non plus seulement avec son ouïe, mais avec son corps et ses gestes... (tout en respectant toutes les règles de bienséance en public).

Composer

Le compositeur n'est pas détrôné non plus, puisqu'il peut toujours, si c'est son désir, composer (poser ensemble) les notes sur une partition et les faire jouer, mais il peut de plus composer des instruments, composer des instrumentistes, et enfin, grâce à cette objectivation nouvelle, celle de "l'objet gestuel" capté, enregistré, représenté, composer le geste.

