OBJETS SONORES, PHÉNOMÉNOLOGIE ET SCIENCES COGNITIVES
Questions sur la modélisation de la perception auditive et musicale

André Villa
CICM – Université de Paris 8
MSH Paris Nord
villaandre@hotmail.com

RESUME

On envisage dans cet essai de tisser quelques lignes interdisciplinaires afin d'analyser le fonctionnement de certains aspects des structures de notre perception auditive. Le but est d'aborder et de développer certaines questions fondamentales à la modélisation informatique de la perception auditive et musicale.

L'idée conductrice de la recherche est de diviser les structures de la perception en trois niveaux qui se « *corrélationnent* » par des projections d'isomorphismes structurels. Ensuite, notre recherche vise à appliquer l'approche la plus appropriée à la modélisation de chacun des niveaux. Finalement nous abordons certains essais qui nous ont déjà présenté les possibilités d'interconnexion (voire *corrélation*) entre les modèles.

Les trois niveaux en question sont-ils : 1) le niveau des objets physiques « dans le monde » (qui sont les « données » de notre perception du monde extérieur), 2) le niveau des structures appelées de *bas niveau* (des systèmes sensoriels jusqu'aux cortex primaires), et 3) le niveau *cognitif* qui, « appliqué » aux « données » du monde extérieur, construit - saisit leurs significations pour la conscience percevante.

On finit par vérifier les contraintes et les possibilités d'une modélisation informatique mixte de la perception auditive et musicale. Notre recherche aborde principalement les niveaux 2 et 3 supra cités. On propose une approche dynamique connexionniste (subsymbolique) appliquée au *bas niveau* et une approche computationnelle (symbolique) appliquée au niveau *cognitif*.

1. INTRODUCTION

Dans cette intervention, nous allons vérifier quelques points du fonctionnement de la perception auditive et des **« représentations mentales » des objets sonores**[1]. Entre les plusieurs points de ce complexe sujet, nous allons prioriser les questions suivantes. Comment pouvons-nous modéliser la représentation de la morphologie d'un objet dynamique comme l'objet sonore ? Comment les ondes de pression acoustiques encodées par le système auditif périphérique peuvent êtres décodés en structures porteuses (en puissance) de signification et reconnaissables par la conscience percevante ? Comment pouvons-nous essayer de modéliser les relations d'isomorphisme entre les objets sonores « dans le monde », les structures internes de l'encodage neuronal du cortex auditif et les représentations mentales du son perçu ? Autrement dit, comment peut-on modéliser le traitement des signaux sonores dans notre perception ?

Quelques questions analogues à celles-ci sont au cœur des plusieurs essais récents de naturalisation de la phénoménologie de la perception[2]. Bien évidemment, l'espace qui nous allons consacrer ici aux possibles réponses est assez restreint et nous oblige à rester à la superficie du problème. Néanmoins, nous croyons que cela n'empêche pas que cette analyse puisse avoir une valeur heuristique pour les recherches en informatique appliquée à la perception auditive et musicale. Nous allons donc, dans un premier moment, orienter cet essai vers une analyse philosophique de la problématique. Ensuite nous allons aborder certains essais récents de modélisation informatique de la perception musicale.

À titre méthodologique, nous allons utiliser quatre approches interpolées : la phénoménologie, la pragmatique, les théories neo-gestaltiques et les données expérimentales issues des sciences cognitives. Nous allons séparer (et interconnecter) trois niveaux de relations entre les structures qui caractérisent le phénomène de la perception des objets sonores. 1) Les structures du percept dans le « monde réel », c'est-à-dire, la partie physique du phénomène. Celle-ci nous est accessible par des rapports de continuité / discontinuité, stabilité / instabilité, singularité / multiplicité. L'objet qui se donne à la perception est à ce niveau composé, par exemple, par des ondes de pression d'air. L'objectif ici est de donner des indices pour une future investigation sur comment arrivons-nous à extraire et à percevoir des discontinuités, voir des singularités dans l'océan de phénomènes physiques qui baignent nos systèmes

[1] La notion d'objet sonore utilisé ici est basée sur le concept proposé par Pierre Schaeffer [22]. C'est-à-dire, l'objet sonore est le corrélat de l'acte perceptif qui le constitue : l'*écoute réduite*. Selon Schaeffer, l'écoute est une *activité* qui peut se tourner vers une des trois possibilités de perceptions extérieures : 1) la perception de l'origine du son (les *indices*). 2) La perception du sens du son (ses *valeurs* relatives à un langage sonore déterminé). 3) La perception associée à la pratique de l'*audition acousmatique* – en se débarrassant du complexe audiovisuel qui est l'origine du son l'intention d'écoute est tournée *vers le son lui-même*. Schaeffer appelle cette écoute en situation acousmatique d'*écoute réduite*. L'objet sonore est donc constitué dans l'action de percevoir un percept sonore spécifique dans une situation d'écoute acousmatique [1].

[2] Cf. par exemple les divers essais dans les ouvrages « *Naturaliser la phénoménologie* » (op. cit. in [27]) et « *Philosophies de la perception* » (op. cit. in [21]). Cf. aussi l'ouvrage « *Physique du Sens* » [20].

sensoriels. Ce niveau, dans la limite de ce texte[3], ne sera pas traité de façon approfondie. 2) Le niveau d'encodage et le traitement du signal (décodage) à *bas niveau*. Nous ferons une analyse des certains aspects des structures des systèmes sensoriels mais principalement au niveau du décodage dans le cortex primaire (e.g. A1 pour l'audition ou V1 pour la vision). Nous allons aborder quelques évidences d'une relation d'isomorphisme psychoneuronal à ce niveau (i.e. la similarité purement structurelle entre les morphologies de l'objet-percept et l'objet qui émerge de l'encodage dans la structure neuro-corticalle) et leurs implications dans le traitement des donnés au niveau cognitif. 3) Le niveau des processus cognitifs. Ce processus à *haut niveau* « construit et saisit la signification » en décodant le sens des structures « émergentes » des cortex sensoriels primaires impliquées dans le traitement des informations pendant le processus de la perception des objets sonores. Notre but est de vérifier les possibilités et les impossibilités (théoriques et empiriques) d'application d'une méthode qui combine les approches subsymboliques et symboliques pour les modélisations de la perception auditive et musicale.

2. LES BASES PHÉNOMÉNOLOGIQUES

Dans les *Recherches Logiques* [9], Husserl travaille et développe le concept de *remplissement* d'une signification. Si je dis : « la table rouge », je n'ai pas besoin de voir une table rouge pour comprendre la signification de la phrase. La signification n'est pas un objet dans le monde, elle est immanente à la proposition. Mais si je vois la table rouge en question, il y a un *remplissement* . C'est ce *remplissement* qui m'assure qu'un être réel correspond bien à cette « visée » qu'est la signification. Cette vérification des propositions et ce remplissage deviennent possibles grâce à une mise en présence de l'objet. Tandis que si un objet « du monde extérieur » est donné à ma perception, je ne « le » retrouve pas dans ma conscience. Il n'est pas un objet mental, immanent à ma conscience. Cependant il est indissociable des impressions contenues dans ma conscience. La conscience « vise » la signification de l'objet aperçu. Husserl appelle cette « visée » d'*intentionnalité* et l'objet de cette visée d'*objet intentionnel*.

L'*objet intentionnel* tout comme la signification **n'appartient pas au monde extérieur**. Évidemment, je peux **imaginer** quelque chose qui n'existe pas dans le monde extérieur. Je peux ainsi avoir la même structure d'objet intentionnel soit visé à travers des contenus de sensation (de voir une « table rouge », objet du monde extérieur), soit visé dans l'*acte* d'imaginer une « table rouge ». L'*acte de conscience* a donc deux dimensions qui sont : 1) l'objet intentionnel contenu en lui-même, lequel nous fournit la « matière » de l'acte ; 2) l'acte lui-même que vise cette matière. Cet acte lui-même peut être une perception mais aussi une imagination un jugement, un souvenir, etc. On peut donc faire une **analogie heuristique** dans un sens de naturalisation de la phénoménologie entre 1) l'objet intentionnel et les *patterns* d'activations neuronales à *bas niveau* de notre cortex, et entre 2) l'acte lui-même et les fonctions cognitives, voir les activations à *haut niveau* de notre cerveau.

2.1. Choses et vécus : les deux perceptions.

Une des distinctions essentielles de la phénoménologie est basée sur la différence des modes d'apparition de l'objet pour l'esprit percevant. Pour Husserl, le dualisme *pensé - choses matérielles* ne s'agit plus de différences entre deux substances fermées sur elles-mêmes (le dualisme cartésien) mais de deux façons d'exister : « *être comme vécu et d'être comme chose* » ! [7]. Il y a donc pour Husserl **deux modes de perceptions**. L'une *immanente* qui perçoit les vécus, et l'autre *transcendante* qui perçoit les choses. En s'opposant, elles distinguent les deux régions auxquelles elles se rapportent. Le *vécu* est *immanent* à la conscience car il n'y a rien en lui qui soit hors de la conscience qui le vise et, dans ce sens, il est « *indubitable* ». Tandis que la chose qui est l'objet de la perception se présente à nous par esquisses : elle a toujours une face cachée, une note qui était masquée, etc. Telle *chose* est par essence *transcendante*, puisqu'elle n'est jamais intégralement accessible en une seule fois à la conscience. Elle est donc, dans ce sens, « *douteuse* ».

Cette distinction est aussi une caractéristique essentielle de la méthode propre à la phénoménologique proposée par Husserl. Il ne s'agit pas de définir la constitution de l'objet *en-soi* – comme le ferait un acousticien pour les signaux sonores – mais de **décrire les structures qui caractérisent la manière dont les objets se présentent**. En autres mots, la phénoménologie analyse les structures qui font le lien entre les deux instances fondamentales, voir les deux façons d'exister, d'un même phénomène. Ce phénomène qui apparaît comme la corrélation entre une intention (la *noèse*, à la fois un acte donateur de *sens* et une visée à vide d'un objet) et le corrélat de cette intention (le *noème*, à la fois subordonné à la *noèse* mais indépendant car il est l'unité objective des déterminations). Cependant les phénomènes ne « parlent » qu'une fois transformés en objets. Cette objectivation étant une construction, constituée à partir d'un donné comme objet, dans la corrélation *noème-noèse*. Comme nous fait remarquer Jean-Luc Marion – même si à l'évidence Husserl n'était pas tout à fait d'accord – l'acquis fondamental de la phénoménologie husserlienne est : **phénomène (*Erscheinung*) ne se dit pas d'abord, ni seulement, de l'objet qui apparaît, mais bien du vécu dans lequel et selon lequel il apparaît** [16].

2.2. La phénoménologie de la conscience du temps.

[3] Au cours du texte nous feront quelques citations des théories impliquées dans les analyses de ce niveau 1. Cependant pour avoir un abordage ponctuel dans ce domaine, cf. [3], [18], [20], [21], [24] et [25].

Husserl a placé la temporalité comme un axe fondationnel et au centre de ses recherches en phénoménologie. Toutes les autres formes d'activités mentales en dépendent. Nous allons utiliser quelques-unes de ses remarques pour analyser deux aspects fondamentaux de notre recherche. 1) La question de l'unité de l'objet sonore et 2) la modélisation connexionniste de la perception de l'objet sonore à *bas niveau*.

Chez Husserl, la question de la « conscience du temps » ne s'agit pas (seulement) de la conscience du temps lui-même ou du fait que la conscience a lieu dans le temps. L'aspect qui l'intéresse est la conscience des objets temporels en tant que temporels. Évidemment une table, une maison ou une mélodie sont des objets qui ont une durée et donc temporels. Mais la conscience du temps qui nous intéresse ici (phénoménologique) est une conscience d'un objet temporel spécifique (l'objet sonore) pendant sa durée simultanément physique et psychique. Autrement dit, pendant le processus de perception du phénomène qui constitue l'« objet sonore », le corrélat d'une écoute orientée et spécifique : l'*écoute réduite* schaefferienne [22]. Cela implique qu'il soit aussi une structure qui transcende le flux des perceptions (i.e. la multiplicité des instants de la perception d'un objet dynamique comme un son) de l'écoute réduite dans son propre flux (i.e. le flux *continuum* de l'acte perceptif). Autrement dit, l'objet sonore consiste dans la permanence d'une forme – la forme du *Jetzt* – associée à une matière qui se remplace continuellement et continûment. Cela c'est la propre définition de la notion de durée appliquée aux vécus chez Husserl [2]. Ce qui dure (de forme intentionnelle) dans la perception auditive ne peut être rien d'autre que l'objet sonore. La matière qui se remplace ne peut être que la corrélation entre les changements de l'onde de pression d'air, dans le monde réel, et les instants ponctuels (discrets) qui forment le flux des impressions, la « multiplicité continue », dans notre perception. Il faut remarquer qu'on peut essayer de modéliser cette notion d'instants ponctuels dans un système dynamique non linéaire subsymbolique en faisant appel, par exemple, à la théorie morphodynamique [20].

Husserl nous fait les remarques suivantes. « *Que plusieurs sons successifs donnent une mélodie, c'est possible seulement parce que la succession des processus psychiques s'unifie « sans plus » pour former un ensemble. Ils sont l'un après l'autre dans la conscience, mais ils tombent à l'intérieur d'un seul et même acte d'ensemble* » [8]. Le philosophe fait encore la distinction entre le « son qui dure » (remplacé dans notre recherche par l'expression *objet sonore*) et « phénomène du son » (le flux des perceptions de notre *écoute réduite*). Il cite : le son que j'entends, « *il lui appartient le « maintenant », le temps phénoménologique, la durée, etc. Distinction en outre entre le son qui dure et le « phénomène du son » comme flux du temps dans lequel le son est identiquement le même* ». Cette distinction se fait nécessaire car l'objet sonore dans le monde est un donné par opposition à la *représentation symbolique*[4] qui le vise et le construit mais qui ne le « donne » pas. Cette construction se fait à partir d'une donnée. Cette donnée est la donnée de la discipline phénoménologique et donc, on a déjà dit, une donnée phénoménale. Cependant Husserl, en évitant possibles interprétations dualistes de ces distinctions, nous réaffirme : « *nous ne prenons pas les contenus que nous vivons comme signes pour autre chose mais nous les prenons tels que nous les trouvons devant nous* » [11].

2.3 La structure tripartite de la temporalité.

En faisant une analyse de la perception des objets temporels, Husserl a développé la notion de la « structure tripartite de la temporalité » : ***rétention, impression originaire et protention***. Par exemple, quand on écoute une mélodie, on n'a pas besoin d'attendre qu'elle soit finie pour la percevoir comme une mélodie. Pendant qu'on écoute une note au milieu de la mélodie, on garde dans l'esprit les notes que viennent toute juste de passer. Si ce n'était pas le cas, on entendrait note par note de forme isolée et pas la mélodie comme une structure – une *Gestalt* – qui forme une unité. Ces notes qui restent dans l'esprit c'est ce qu'Husserl appelle de ***rétention***. Ce n'est pas du souvenir (ou de la mémoire) car ce sont des impressions qui sont **encore présentées à l'esprit, même que sur la forme d'absence**. La rétention a lieu « maintenant », pendant la perception de la mélodie. On n'a pas besoin d'y aller chercher ces donnés ailleurs dans notre esprit comme dans le cas d'une « *working memory* ». La rétention, différemment du souvenir, n'est pas objectivante (souvenir en tant qui réflexion immédiate sur ce qui vient d'être vécu). Le souvenir, Husserl nous le fait remarquer, différemment de la rétention, est une saisie « en réflexion » [7], [9] et [10].

Dans un autre sens, Husserl nous montre qu'on a aussi une « chose intuitivement attendue » et qu'en on a conscience par anticipation comme « à venir ». Par exemple, dans le même cas de la note au milieu de la mélodie, pour les notes qui vont arriver. Cela est valide même dans le cas d'une mélodie inconnue. C'est ce qu'il appelle de ***protention***. La protention est fondamentalement présentée par Husserl comme une sorte d'opposée symétrique de la rétention, dirigée plutôt vers le futur que vers le passé. Bien évidemment, différemment de la rétention, la protection n'a pas encore eu lieu et ce future est encore ouvert. D'ailleurs c'est quelque part aussi avec cette attente que les compositeurs travaillent leurs compositions vis-à-vis de l'auditeur. Parfois casser l'attente avec l'inespéré est une bonne forme de rendre une œuvre « créative », « originelle ». Parfois rendre l'espéré est une bonne forme d'affirmer, voir confirmer, la qualité d'une écriture (e.g. un accord de septième de dominante – de la

[4] L'expression *représentation symbolique* est l'expression utilisée par Husserl [11]. Dans notre recherche, elle fonctionne comme un synonyme d'*écoute réduite*. On a préféré la garder car elle invoque déjà l'approche computationnelle abordée plus bas.

sensible – a la fonction de générer une tension et produire chez l'auditeur une "attente" de repos, ou une gamme « correctement » appliquée confirme un certain style d'harmonisation ou une certaine cadence, etc.). Cependant la question que si pose ici n'est pas d'ordre esthétique mais plutôt : comment est-il possible que la protention soit d'une certaine façon à la fois déterminée et ouverte ? On y reviendra.

Le troisième terme de la structure de la temporalité est l'***impression originaire***. C'est l'instant où le percept qui s'imprime dans nous systèmes sensoriels si met en corrélation avec l'intentionnalité qui le vise. Évidemment cet instant est impossible d'être déterminé précisément. Car, quand est-ce que, par exemple, un son s'imprime-t-il dans notre système auditif ? Quand est-il filtré par le pavillon ou quand est-il amplifié par le canal de conduite auditif externe ? Quand touche-t-il le tympan ou quand est-il encodé et transformé en potentiels d'action (*spikes*) par la cochlée? Cependant il ne faut surtout pas oublier que l'impression originaire n'est pas simplement une impression causée par les stimuli sur nos systèmes sensoriels. Or, la perception, comme l'on a déjà dit, est une action où l'attention dispensée par le sujet qui perçoit, qui jette son *intentionnalité* sur le *Dasein* (le *être-là* dans le monde). Cette action est également la cause du phénomène. Quand est-ce donc que cette intentionnalité « atteint » son objet visé ? Pour instant, nous allons juste retenir que l'impression originaire se présente plutôt indéterminée, comme le début de l'amalgame entre intérieur et extérieur, entre *intentionnalité* et *objet intentionnel*. Cependant elle se présente toujours comme l'articulation entre la rétention et la protention dans l'activité de l'esprit percevante. Plus encore, les impressions originaires sont à l'origine des rétentions.

Basé sur ces analyses, Husserl nous montre que le présent (« maintenant ») n'est pas une « *lame d'un couteau* » qui sépare le future du passé. Le présent est quelque chose de plus large temporellement et qui enveloppe la structure tripartite ***rétention, impression originaire et protention***. Cette structure tripartite forme une « multiplicité continue ». Il faut remarquer que la rétention et la protention ne s'étendent pas infiniment dans le passé et le futur. Quand on entend une œuvre musicale, on n'a pas forcément à l'esprit la première et la dernière note de la pièce. Le concept de présent dans ce sens phénoménologique est pris comme un « *présent spécieux* ».

Cette notion est devenue un des incontestables arguments en faveur de la modélisation connexionniste (au moins à bas niveau) de la perception des objets sonores [27] et [29]. Par exemple, Van Gelger, qui utilise dans ses recherches le modèle LEXIN (figure 1) des systèmes dynamiques de reconnaissance de *patterns* auditifs, nous fait remarquer que[5] : 1) une interprétation dynamique de la rétention l'identifie à un aspect de l'état actuel du système, plutôt qu'à n'importe quel aspecte passé ou futur. 2) La localisation de l'état actuel du système reflète non seulement les sons antérieurs auxquels il a été exposé, mais également l'ordre dans lequel il l'a été – ou, plus généralement, depuis combien de temps, il l'a été. Par conséquent, si l'on suit cette interprétation, cela a un sens clair de dire que la rétention se rapporte au passé en tant que passé à un certain degré. 3) Une impression originaire est ce « moment » de la conscience du temps qui se rapporte à la phase actuelle de l'objet temporel. Dans le cas de la reconnaissance d'une mélodie, c'est l'aspect de notre conscience de la mélodie comme un tout qui est la conscience de la note actuelle. Pour l'explication dynamique, la conscience de la mélodie comme un tout est identifiée à l'état actuel du système de reconnaissance auditive. 4) La protention est identifiée à l'occupation par le système de sa localisation actuelle, puisque c'est cette localisation (associée à la dynamique intrinsèque du système, et si les influences extérieures sont maintenues constantes) qui détermine le comportement futur du système. Comme pour la rétention, remarque encore le chercheur, cette explication donne un sens à l'idée que la protention est actuelle, qu'elle se rapporte aux périodes futures en tant qu'elles sont futures à un certain degré, qu'elle est une « multiplicité » continue, qu'elle est finie et directe.

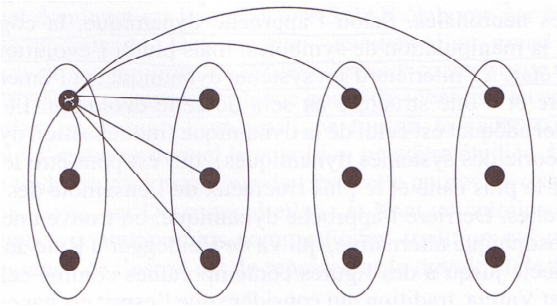


Figure 1 – Une représentation schématique de l'architecture du réseau de neurones du modèle dynamique LEXIN de reconnaissance de *patterns* auditifs. Chaque unité inhibe chacune des autres unités à l'intérieur d'un même groupe et excite celles qui appartiennent à d'autres groupes.

Dans le cas d'une modélisation computationnelle du « présent spécieux », le pattern auditif est d'abord encodé symboliquement et stocké dans le *buffer*. C'est seulement après le stockage du *pattern* entier dans cette *mémoire à court terme* que les algorithmes de reconnaissance peuvent être appliqués au traitement des informations. Et c'est seulement après le travail des algorithmes que l'on peut dire qu'un type de conscience du *pattern* en question émerge du système. Le problème c'est que, comme nous fait remarquer Husserl, nous n'attendons pas que la mélodie soit finie pour l'entendre comme une mélodie. Ceci se passe en même temps en que nous la percevons, c'est-à-dire, pendant qu'elle est en cours de diffusion ou exécution.

2.4 L'objet sonore et la singularité phénoménologique

[5] Nous citons délibérément à ce point certaines remarques très pertinentes faites par Van Gelger [27].

Un des aspects indispensables à la caractérisation de la singularité phénoménologique est l'approche que va prendre le rapport entre l'objet sonore et les structures données à la perception. Tout d'abord parce que la relation objet-structure est à la base de notre activité perceptive. Tout objet n'est perçu comme tel que dans un contexte formé par une (ou plusieurs) structure qui caractérise la « forme » de l'objet (cf. les lois de la *Gestalt*, [6] et [12]). Cette relation est vitale à la définition de l'objet sonore comme l'« unité ». L'objet sonore nous apparaît donc comme une *Gestalt*, un ensemble formant une structure perçue comme une unité[6]. Cependant cette *Gestalt* est aussi une structure en tant qu'un « composé de plusieurs objets ».

Ainsi tout objet est **identifié** comme objet dans une structure. Cette structure peut à son tour être **identifiée** comme objet dans une autre structure originelle, formée par d'autres objets constituants. On remonte ainsi vers l'infiniment grand, vers une échelle **macro** (e.g. une mélodie – ou un motif – peut se présenter comme une *Gestalt* dans une structure comme une phrase musicale. Une phrase peut se présenter comme une *Gestalt* dans une autre structure, à un autre niveau, comme dans une œuvre musicale, etc.). En contrepartie, toute structure n'est perçue que comme structure d'objets composants et on peut donc **qualifier** un objet de cette structure comme structure originelle d'autres objets constituants, et ainsi de suite. On descend ainsi vers l'infiniment petit, vers une échelle **micro** (e.g. cette même mélodie est formée par d'autres structures – les notes – qui peuvent à la fois être prises comme une structure formée par des structures tout aussi composables – les harmoniques, les partiels inharmoniques, l'attaque, l'entretien, l'extinction, etc.). Cette perspective *multi-échelle* des objets sonores est devenue très évidente dans le domaine de la composition assistée par ordinateur. Comme l'a remarqué Horacio Vaggione, « *les techniques d'analyse, synthèse et transformation du son offrent la possibilité de générer des morphologies projetées dans un vaste réseau concernant une articulation de ces multiples niveaux* » [26]. Désormais le compositeur contemporain doit travailler les corrélations entre les singularités et les multiplicités qui peuvent émerger d'un même complexe sonore de plus en plus plural. Ces rapports *multi-échelle* génèrent des nouveaux espaces composables et tout un monde de possibilités à être exploré, tantôt par le compositeur que par les auditeurs.

Mais on utilise donc le mot structure dans deux sens distincts. Selon l'approche de d'identification (structure comme contexte où l'on identifie un objet) ou de qualification d'un objet (structure comme *Gestalt*, forme, organisation propre à l'objet). Par rapport à ceci, une question qui s'impose concerne l'ambiguïté ontologique des structures. Umberto Eco nous interpelle : « *la structure est-ce un objet en tant qu'il est structuré, ou bien l'ensemble des relations qui structurent l'objet mais que l'on peut abstraire de l'objet ?* » [5]. Or, basé sur les concepts phénoménologiques supra cités, on peut affirmer que la structure (comme l'essence, voir l'*eidos* de l'objet) n'est pas détachable de la substance où elle s'actualise. Elle est à la fois, nous dirait Jean Petitot, « *ossature intelligible et objet structuré* » [18]. Mais elle n'est ni simplement donnée par leur composition physique, ni simplement construite par l'esprit qui la perçoit : elle émerge au carrefour de la corrélation *noème-noèse*. Il faut donc comprendre les verbes *identifier* et *qualifier* dans cette relation.

L'autre question qui nous intéresse ici c'est l'émergence des singularités des systèmes naturels et leurs relations avec la double échelle macro/micro. Car, depuis quelques années, les progrès en physique et mathématique nous ont beaucoup aidé dans la compréhension des phénomènes d'(auto)organisation des substrats matériels (la théorie des singularités, la théorie qualitative des systèmes dynamiques non linéaires, la théorie des phénomènes critiques, etc.). Comme nous fait encore remarquer Petitot, les systèmes naturels sont des systèmes à (au moins) deux niveaux de réalité objective. Il y a un niveau micro qui correspond à la physique fondamentale du système. D'autre part, il y a un niveau macro, qui est de nature plus *morphologique* – dans le sens de forme émergente qui détermine tout ce qui concerne l'organisation et la structuration qualitative (spatio-temporelle) des formes sensibles et des formes naturelles (sonores, visuelles ou autres) – que physique.

Le niveau macro émerge donc du niveau micro sous-jacent et l'on peut contrôler mathématiquement le processus de changement de niveau dans ces modèles. Ce sont essentiellement les comportements collectifs coordonnés et coopératifs des entités micro locales qui déterminent les morphologies au niveau macro. Le connexionnisme est une approche qui exemplifie ce constat. Le niveau macro, comme nous rappelle le mathématicien, « *est essentiellement organisé autour des singularités (caustiques, transitions de phases, ondes de choc, défauts, ruptures de symétries, etc.) des processus physiques sous-jacents. Ces singularités portent l'information et sont phénoménologiquement dominantes. C'est à travers elles que s'effectue la structuration qualitative, l'organisation morphologique des phénomènes* » [19]. Il nous fait encore remarquer qu'il y a de fortes propriétés d'universalité des comportements critiques. Autrement dit, il y a une notable indépendance de l'organisation du niveau morphologique émergent (macro) par rapport à la structure micro. Cette indépendance est basée sur les règles morpho-structurales des phénomènes relatives de la physique du niveau d'infrastructures catastrophiques[7] (micro), sous-jacente.

[6] Cf. les catégorisations typo-morphologiques des objets sonores du *Traité des objets sonores* de Schaeffer [22]. Ces catégorisations sont basées en grande partie sur les rapports d'intensité relative des accidents du profil de l'enveloppe amplitude / temps de l'objet sonore.

[7] Pour la théorie des catastrophes, cf. René Thom [24] et [25]. C'est une théorie qui s'efforce de décrire les discontinuités qui peuvent se présenter dans l'évolution d'un système. « *Intuitivement, on admet que l'évolution globale d'un système se présente comme une succession d'évolutions continues, séparées par des sauts brusques de nature*

3. L'ISOMORPHISME

La notion d'isomorphisme qui nous utilisons ici est un mélange des concepts de la chimie, de la mathématique, de la philosophie et, bien entendu, de la *Gestalttheorie*. En chimie, l'isomorphisme est une propriété que possèdent deux ou plusieurs corps d'avoir de formes (cristallines) similaires. En mathématique, l'isomorphisme c'est un morphisme dont l'application est bijective. C'est-à-dire qu'à tout élément de l'ensemble de départ, est associé un et un seul élément de l'ensemble d'arrivée.

En philosophie, une bonne présentation pragmatique de la notion d'isomorphisme est donnée par Wittgenstein. « *Le disque de phonographe, la pensée musicale, la notation musicale, les ondes sonores sont tous, les uns par rapport aux autres, dans la même relation représentative interne que le monde et la langue. À tous est commune la structure logique… Qu'il y ait une règle générale grâce à laquelle le musicien peut extraire la symphonie de la partition, et grâce à laquelle on peut extraire la symphonie des sillons du disque, et derechef, selon la première règle, retrouver la partition, c'est en cela que repose la similitude interne de ces figurations apparemment si différentes. Et cette règle est la loi de projection qui projette la symphonie dans la langue de la notation musicale. C'est la règle de traduction de la langue de la notation musicale dans la langue du disque* » [31]. On pourrait même ajouter que les représentations des ondes de pression d'air – entre autres – qui sont monnaie courante en informatique musicale sont aussi des formes de projection isomorphe du son qui y est représenté.

3.1 Isomorphisme psychoneural.

L'hypothèse de l'isomorphisme psychophysique ou psychoneural proposée par le psychologue gestaltiste Wolfgang Köhler suppose une correspondance entre les représentations mentales d'un phénomène perçu et les états d'activation du cerveau ou des représentations corticales.

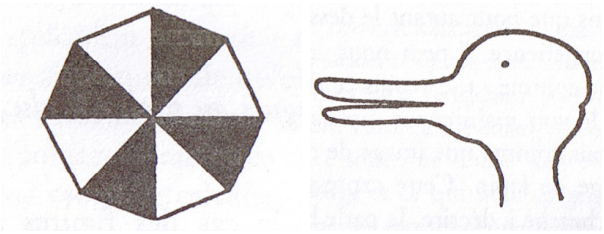
L'isomorphisme gestaltiste n'implique cependant aucune identité morphologique simple. On ne retrouve pas une représentation identique de la carte de la France dans nos aires corticales visuelles si jamais on regarde telle carte. Il faut comprendre l'isomorphisme psychoneural comme une projection structurelle, topologique, dynamique, géométrique mais aux sens abstraits et multiples que les mathématiques donnent à ces termes. Cependant l'isomorphisme ne signifie pas non plus une sorte d'implantation ou d'inscription de structures mathématiques dans notre cerveau. Son support, le tissu nerveux, lui impose des contraintes physiques spécifiques. Mais les formes qui s'y présentent restent susceptibles de transpositions car elles sont faites de relations physiques dynamiques susceptibles de se réaliser en différents champs. Le concept d'isomorphisme, on l'a fait remarquer, n'est à la base qu'une dérivation d'un concept de « transposabilité » par projections structurelles basées sur des règles de traduction des projections concernées. Néanmoins les psychologues gestaltistes n'avaient pas à l'époque les connaissances empiriques des structures corticales qui nous sont maintenant accessibles, ni les récents progrès de la mathématique et de la physique. Bien entendu, cette hypothèse de l'isomorphisme psychoneural a été durement attaquée par quelques critiques.

Après 1947, date où paraît l'édition définitive de *Gestalt Psychology* de Köhler [12], Wittgenstein lui consacre plusieurs séances de son *Cours sur la philosophie de la psychologie* et le premier volume des *Remarques sur la philosophie de la psychologie* [30] et [32]. Une des critiques faites à l'isomorphisme psychoneural se base sur les remarques suivantes. On peut parler d'une *table*, pas seulement quelle est une *table*, mais aussi parce qu'on peut la « *voir comme* » ou se « *souvenir de ça comme une table*». On est donc tenté supposer qu'on doit à plusieurs fois avoir quelque sorte de représentation interne qu'est en soi-même une *table*. Mais certains critiques, et notamment Wittgenstein, ont montré qu'on apprend l'usage approprié de tels mots comme « *table* » et leurs significations d'une communauté verbale qu'est accessible seulement pour l'objet public et pas pour telles images privées. Le philosophe soutient qu'il n'est pas nécessaire, pour comprendre ou expliquer un concept, de pouvoir en donner une définition caractère par caractère, ce qui impliquerait que le concept en question ait des bords nets. Ces notions, que Wittgenstein développe dans le concept de « air de famille », sont applicables aux concepts philosophiques les plus massifs (les « concepts formels ») tels que proposition, langage, nom, nombre mais aussi aux concepts psychologiques.

L'autre notion développée par Wittgenstein et impliquée dans ces analyses est la notion d'« *aspect* ». Les exemples typiques sont ceux de la *croix noire et blanche* et de la figure du *canard-lapin*. Dans les deux cas, on *regarde* le même stimulus (figures 2a et 2b). Cependant pour, par exemple, la figure 2b, on la *voit comme* un canard et on la *voit comme* un lapin. Il existe donc des *aspects* de l'objet perçu qui sont **déterminés par des pensées et des associations**.

---

qualitativement différente. Pour chaque type d'évolution continue, subsiste, en principe, une modélisation de type différentiel classique : mais les sauts font en sorte que l'on passe d'un système différentiel à un autre. La donnée de la théorie des catastrophes apparaît alors comme une sorte de « paquet » de systèmes différentiels qui sont en nombre fini dans la meilleure des hypothèses. Donc, le point représentatif « saute » d'une évolution continue décrite par un système d'équations différentielles à une autre évolution continue décrite par un autre système et, dans certains cas, on ne peut même pas exclure qu'un nombre fini de systèmes ne soit pas suffisant pour décrire la situation tout entière. »



(a) (b)

Figure 2 – (a) La croix noire et blanche et (b) le canard-lapin [30]. Notre pensée peut faire basculer entre les *aspects* des images. Contrairement à ce qui proposé Köhler, une partie des projections isomorphiques n'est pas *innée* mais déterminée par notre pensée et par conséquent .

Mais il existe aussi d'autres *aspects* qui sont « **purement optiques** ». Par exemple, nous savons aujourd'hui que le plus simple des principes gestaltistes (la loi de la *bonne continuation*) est, fort probablement, dû aux contraintes biologiques de notre aire visuelle primaire (V1). Plus précisément aux structures colomnaires et hypercolomnaires de la sous-couche 4C. Ainsi, par exemple, dans l'illusion du triangle de Kanizsa (figure 3), les « bordes illusoires » et la brillance du triangle blanc ne sont pas fruits d'un traitement cognitif (*top-down*). Ce n'est pas le fait que nous avons la connaissance de la forme et du concept de « triangle » qui nous fait inférer cette illusion de la figure. C'est le fait que nos neurones en 4C, responsables pour les représentations rétinotopiques des positions et d'orientations de la borde imaginaire, **sont allumés**. Notre cortex primaire « voit » le triangle sans le soumettre à notre pensée. On ne peut pas éviter cette illusion volontairement (si l'on regarde la figure, bien évidemment).

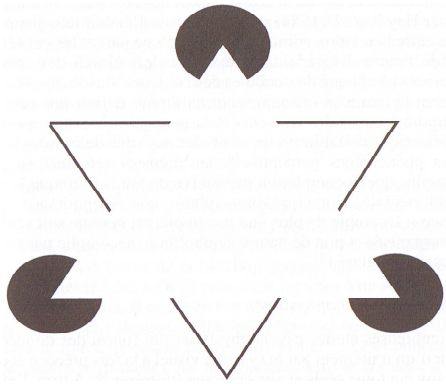


Figure 3 – Le triangle de Kanizsa. Les récentes théories mathématiques appliquées à la neurobiologie nous montrent que cette illusion est, fort probablement, due majoritairement aux structures neuro-corticales de V1. On « voit » les bordes grâce à l'activation des neurones qu'y sont concernés dans le champ récepteur visuel. Différemment du *canard-lapin*, si l'on regard attentivement la figure, on ne peut pas éviter l'illusion ou la faire basculer dans une *non-illusion.*

Ce que Wittgenstein reproche à Köhler c'est le fait que l'*aspect*, du moins jusqu'à un certain point, est soumis à la volonté. Autrement dit, l'isomorphisme psychoneural ne dépend pas seulement des structures des objets et ses représentations « *innées* » (voir projections structurelles isomorphiques) corticales à bas niveau (*bottom-up*), mais aussi des activités et des traitements cognitifs (*top-down*). Plus encore, « *les comptes rendus d'apparition d'aspect ne sont pas des descriptions d'une expérience, intérieure, mais des réactions spontanées à ce que nous voyons, des extériorisations. De plus, ce qui change dans l'apparition d'un aspect n'est pas ce que nous percevons ou son organisation, mais plutôt notre façon d'y réagir, notre attitude envers le vu, et ce que nous allons faire avec lui, notamment dans le cas de la perception esthétique (nous changeons soudain de manière de jouer, de tempo par exemple, quand nous remarquons dans le morceau de musique un aspect jusque-là inaperçu* » [4]. Un autre exemple auquel s'applique la notion de changement d'*aspects* est sur les transformations des paradigmes scientifiques. Comme dans l'expérience où Aristote ne voyait qu'une pierre qui balance (soutenue par un fil), Galilée a vu un pendule. Tous les deux ont vu le même « pendule », mais ils ont donné des interprétations différentes [14]. Pour exemplifier dans le cas des « paradigmes musicaux », les sons qu'il y a 50 ans ne s'entendaient que *comme du bruit* sont aujourd'hui utilisés de façon courante dans les compositions électroacoustiques.

Il nous faut encore remarquer que les questions des figures bistables (comme le canard-lapin ou la croix noire et blanche) sont aussi abordées par Varela. Il nous montre que l'approche de la perception du « *présent spécieux* » par des types spécifiques de dynamiques – comme les groupes d'oscillateurs couplés – sont intéressants car leur comportement n'est généralement pas en accord avec la notion classique de stabilité, dérivée d'une représentation mécanique du monde. À l'opposé, cite Varela [29], dans les systèmes dynamiques comme dans les systèmes biologiques, « *l'instabilité se présente comme la base du fonctionnement* normal, *plutôt que comme une perturbation qui aurait besoin d'être compensée* ». Encore un motif que nous fait opter pour la modélisation dynamique à *bas niveau.*

Dans une lecture phénoménologique de la question, on peut d'abord se demander comment peut-il y avoir des projections de structures d'isomorphisme neuronales dans le cas d'une signification sans remplissage (sans la présence « réelle » de l'objet) ? Par exemple dans le cas d'une pensée musicale, comme dans l'acte de composer (voir « imaginer ») ou de se souvenir d'une mélodie sans l'écouter. Ce type de relation a été compris par certains auteurs comme une sorte d'« isomorphisme de second ordre » [23]. C'est-à-dire, non pas la relation d'isomorphisme entre un objet individuel et les représentations internes correspondantes, mais entre les relations parmi des objets externes alternatifs et les relations parmi ses représentations internes correspondantes. Mais, comme l'on a constaté, ceci a été déjà remarqué par Wittgenstein quelques années auparavant.

Dans cette même méthode phénoménologique, l'on peut dire que les structures de l'isomorphisme de notre perception « représentent » – le mot pris dans son

étymologie originale, c'est-à-dire, *rendre présent* – l'« objet réel » (dans le monde) perçu à la même mesure que l'objet réel perçu « représente » les structures de notre perception. Autrement dit, on peut interpréter, de forme heuristique, le *noème* et la *noèse* comme deux sortes de projections isomorphiques qui, en corrélation, *rendre présent* le même phénomène. L'isomorphisme, l'on a dit, a une acception de similarité purement structurelle entre les choses concernées. Ce sont des projections de structures (formes, *patterns*, *Gestalten*), qui se renvoi les unes aux autres dans des relations de transposabilité (pas forcément bijectives à tous les niveaux) à « double sens » : intérieur vers extérieur et vice-versa, à la fois *bottom-up* et *top-down*.

La question de l'isomorphisme psychoneural à *bas niveau* a été vérifiée dans la structure **physique du cortex** pour quelques recherches. Notamment les découvertes faites sur les projections rétinotopiques. Ce sont des projections neurophysiologiques de la rétine sur les couches corticales. Ces projections sont des applications qui préservent la topographie rétinienne. Il y a donc une représentation conforme logarithmique entre la rétine et la couche 4C de V1 où se projettent majoritairement les fibres issues du corps genouillé latéral (figure 4).

4. SIMULATIONS INFORMATIQUES DE L'ISOMORPHISME PSYCHONEURAL

Au niveau du cortex auditif, cette vérification de projection topographique est beaucoup plus compliquée due à la nature dynamique des objets sonores. Les recherches dans ce domaine ont, pour l'instant, plutôt trouvé des évidences de l'impressionnante plasticité des aires corticales auditives

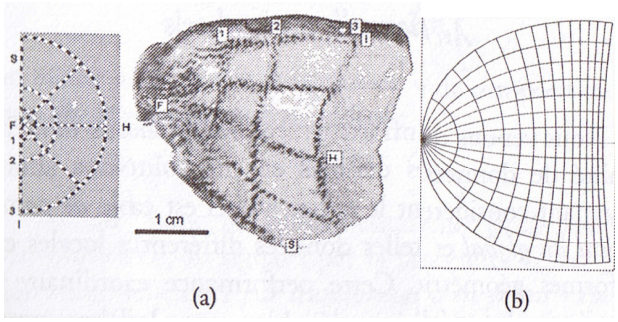


Figure 4 – (a) La projection rétinotopique de la rétine sur la couche 4C de l'aire V1. (b) Modèle mathématique au moyen d'une représentation conforme de type logarithme d'une homographie [21].

que projections topographiques de la cochlée sur A1. Cependant les essais récents de modélisation de la perception musicale par les réseaux de neurones [15] nous présentent quelques résultats de l'hypothèse d'isomorphisme psychoneural à ce niveau (i.e. la similarité purement structurelle entre les morphologies de l'objet-percept dans le monde et l'objet qui émerge de cette **structure neuro-corticalle artificielle**). Car l'hypothèse d'isomorphisme psychoneural peut être soumise à une vérification dans des tests de simulations à l'ordinateur avec une approche de modèle écologique. En d'autres mots, un réseau de neurones artificiels qui par auto-organisation fait émerger un schéma. Un schéma, selon le psychoacousticien Albert Bregman, est une structure de contrôle dans le cerveau humain qui est sensible à certains *patterns* qu'émergent fréquemment, soit dans l'environnement, soit dans notre cerveau ou soit comme interaction des deux [3]. Déjà le chercheur Marc Leman, qui a développé l'expérience qu'on analyse ici, fait la distinction entre l'apprentissage perceptuel d'un schéma *data-driven* à long terme et le contrôle *schema-driven* à court terme. L'apprentissage perceptuel est effectué par des principes dynamiques d'auto-organisation. Les schémas de contrôle sont réalisés par des processus associatifs concernés dans les attracteurs des systèmes dynamiques.

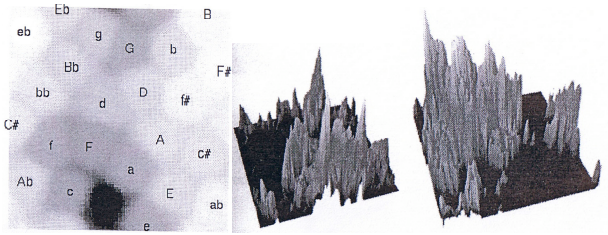
Le paradigme proposé par Leman & Carreras est basé sur l'hypothèse qu'il y a une similarité (isomorphique) entre les activations du réseau de neurones du modèle de l'ordinateur et les représentations neuronales auditives dans notre cortex. Selon cette théorie, quelque part, dans notre cortex auditif, il y doit avoir une aire où les neurones ont une réponse d'activation avec une représentation topographique du cercle des quintes pour la perception des musiques tonales.

L'expérience est basée sur une simulation par ordinateur d'un apprentissage perceptuel des configurations de la hauteur (*pitch,* voire la tonalité) du contexte (une pièce musicale), mais tout ça à des conditions « réalistes ». À savoir, en utilisant de la musique, jouée par un musicien dans un instrument acoustique. La base de la simulation a été un enregistrement (taux d'échantillonnage à 20kHz, 16 bits) du Livre I du *Clavecin bien tempéré*, de J. S. Bach. L'œuvre est composée par 24 préludes et fugues dans toutes les tonalites. Le modèle contient deux parties : 1) un module de perception basée sur une modélisation de l'encodage réalisé par la cochlée (un banc de 20 filtres passe-bande superposés, qui envoie le son transformé in *patterns* de mise à feu neuronales, etc.), et 2) un module de cognition formé par un apprentissage perceptuel type *data-driven* (apprentissage non guidé par aucune connaissance pré-établie, où toutes les informations viennent de l'environnement) et basé sur une architecture Self-Organizing Map (SOM) [13] qui utilise un alignement bi-dimensionel de neurones 100 X 100. Les bords du réseau sont connectés entre eux.

Les schémas de réponses du réseau présentent l'état des activations de tous les neurones du réseau (figure 5a). Ces activations transformées en représentations tridimensionnelles – avec l'addition du niveau d'activation de chaque neurone comme troisième variable – ont donné des ***patterns*** **de représentation 3D très distincts (et donc reconnaissables, par exemple,**

par un algorithme) pour chaque un des 24 tons (figures 5b et 5c).

Les résultats des simulations (reconnaissance par le réseau du ton de morceau joué) ont été comparés avec des tests basés sur des études comportementales appliquées à des sujets humains musiciens (experts et débutants) et non musiciens. Les résultats finaux, bien proches des résultats des sujets humains, ont bien confirmé l'hypothèse originelle des auteurs. Plus encore, ces résultats ont leur incité à développer une autre hypothèse basé sur l'isomorphisme psychoneural. Celle-ci « prédit » que les états des réponses neuronales d'un schéma tonal ont une organisation topographique in termes de cercle des quintes et que cela est la structure de réponse que se trouve aussi dans les études comportementales des humains.



(a) (b) (c)

Figure 5 – (a) Schéma réponse pour le centre tonal dans le réseau neuronal représentant le Do majeur (le noir indique une haute activation, le blanc indique une basse activation. (b) Les niveaux d'activation en 3D du schéma pour Do mineur et (c) Do majeur [15].

4.1 Approches combinant traitements symboliques et subsymboliques.

Le chercheur Udo Mattusch a développé un paradigme qui travaille la combinaison d'un réseau de neurones pour transitions et une description abstraite symbolique de certaines caractéristiques des processus subsymbolique et symbolique. Le but de son approche est l'extension des modèles dynamiques non-linéaires à un traitement linéaire de l'information. Il utilisa pour ça des aspects de la perception de structures (*Gestalt*) musicales que doivent être prises en compte par la combinaison en question [17]. L'auteur travaille avec la théorie que l'invariance des *Gestalten* résulte d'un processus d'assimilation des formes comme base pour l'assertion et la stabilité des figures musicales. Les *Gestalten* possèdent des dispositifs figuratifs que gardent une constance quand certaines transformations sont effectuées. Par exemple, dans la perception des mélodies, certaines formes d'abstraction sont appliquées dans la construction des contours et des intervalles. Cette construction peut être comprise comme un des pas dans le processus d'abstraction de la perception musicale. Elle nous donne, par exemple, la possibilité de la reconnaissance d'une mélodie transformée (e.g. transposé à un autre ton) comme « la même ».

Mattusch a utilisé un système d'opération symbolique de modèles basé sur PROLOG[8], et un système de traitement basé sur un réseau artificiel de neurones (3 couches de rétropropagation) habilité à construire des représentations qui peuvent si relier et ainsi se connecter à chaque une des autres représentations. Pour faire l'entraînement du réseau de neurones, il a utilisé seulement les *patterns* de contour et de l'intervalle de 4 notes. Ces notes sont encodées in langage MIDI et sont associées comme *bit patterns* pour l'encodage du contour et de l'intervalle. Le réseau est composé par trois types de matrices, une IMPUT, une OUTPUT et une TARGET qui contient le vecteur de référence pour lequel le vecteur des donnés de départ du réseau est comparé pendant le processus d'apprentissage.

Le cas de transposition d'une mélodie, selon certains chercheurs, peut être vu comme une variante de la loi de destin commun (*common fate*) de la *Gestalt*. Cette loi postule que des éléments avec les mêmes caractéristiques (e.g. des éléments qui se déplacent ensemble) appartiennent à la même *unité*. Dans l'expérience de Mattusch, les matrices du réseau contiennent en chaque cas la différence de l'intervalle rapporté à un élément. Ainsi, cela permet de grouper facilement tous les éléments ensemble *comme une Gestalt* s'ils sont transformés par le même groupe du réseau neuronal. Une application du système hiérarchique de représentations multicouches à ces configurations permet une commode disposition de structures et substructures parmi plusieurs niveaux de représentations.

Cette approche a, bien entendu, plusieurs limitations. Mais l'auteur nous surligne plusieurs avantages. Par exemple, les possibilités des représentations symboliques comme les inférences logiques issues d'informations manquantes, la possibilité de transférer les fonctions et d'autres aires représentationnelles, ainsi comme la possibilité de connecter ces représentations syntaxiquement et sémantiquement à d'autres types de connaissance. Dans cette approche, tout cela peut être applicable au réseau, gardées certaines restrictions, en faisant l'use d'une correspondance organisée entre les couches du réseaux.

5. CONCLUSION

Nous traitons dans ce texte les possibilités ouvertes par les approches de la modélisation de certains aspects de la perception auditive combinant les méthodes subsymboliques et symboliques. Cette méthode a été déjà proposée par certains auteurs [28], mais à l'époque les applications concrètes faisaient « *encore défaut* ». Cela nous semble être la bonne voie pour la modélisation de la perception auditive et musicale. Mais pas

[8]PROLOG est un programme qui traduit des représentations, comme les configurations des activations d'un réseau de neurones, en prédicats après un apprentissage inductif. Évidemment l'échange d'information entre le réseau neuronal et le système PROLOG demande une adaptation entre les structures de représentation symbolique et les valeurs des matrices.

seulement parce qu'elle nous offre déjà quelques bons résultats. Ces approches nous permettent d'utiliser et de respecter les analyses et les contraintes imposés par plusieurs théories qui se sont occupées de cette question, comme la phénoménologie (la perception des objets temporels, le présent spécieux, etc.), les questions de l'isomorphisme (pragmatique et psychoneural), les lois de la *Gestalt* (de la bonne continuation et du destin commun), les approches neo-gestaltiques (théorie de la morphogénèse, théorie morphodynamique), les théories physiques et mathématiques (théorie des singularités, théorie des catastrophes) entre autres. Sans oublier les possibilités réelles de modélisation par l'informatique.

En résonance avec ces théories, nous proposons donc une approche de modélisation à trois niveaux : 1) la captation des sons et l'encodage par banc de filtres par fenêtrage variable basé sur les recherches des implants cochléaires (sujet à peine cité dans ce texte), 2) le décodage isomorphique subsymbolique avec les modèles connexionnistes de réseaux de neurones, et 3) le traitement computationnel symbolique des *patterns* émergents du niveau subsymbolique.

Le sujet est évidemment beaucoup plus large et complexe que cette infime partie qu'on a abordée dans ce texte. Mais l'interrogation qui émerge ici est la suivante : dans quelle direction pouvons-nous bien mener nos recherches sur la modélisation informatique de la perception auditive et musicale ? Question *sine qua non* pour le développement de l'intelligence artificielle. Certains aspects ici effleurés sont visiblement encore ignorés – ou peut-être regardés avec une certaine négligence – par certaines approches. Notamment les approches purement computationnels (symboliques) dans la modélisation des traitements du langage.

Ces approches, dans l'idée de combler l'écart entre la nécessité d'un niveau sémantique (représentationnel) et la contrainte exigeant que ce niveau soit, après tout, physique, introduisent les symboles à la fois pourvus de signification et physiques. Mais comment, dans cette stratégie, les symboles acquièrent-ils leur signification dans un apprentissage perceptuel du type *data-driven* ? Sans les contraintes de la programmation de l'ordinateur et la prédéfinition de l'ensemble des stimuli, les symboles n'acquièrent pas de signification. Or, la modélisation mixte et à plusieurs niveaux peut, comme on a essayé de démontrer, combler ce déficit. Les recherches de Leman, Carreras et Mattusch ont déjà produit des résultats que nous permettent d'affirmer que la modélisation informatique de la perception auditive et musicale doit se faire dans ce sens.

Cependant, il ne faut pas oublier, notre recherche a pour but aussi la relation entre la perception des objets sonores, l'informatique et, évidemment, la musique. Or, à la différence de Pierre Schaeffer, nous croyions que, avec la numérisation, tous les sons sont devenus passibles d'objectivation. Objectivation dans le sens d'une corrélation qui se donne à la construction par – et avec – une écoute privilégiée, orientée et, pourquoi pas, *réduite*. Plus encore, tous les objets sonores sont porteurs, en puissance, de formes musicales, voire d'objets musicaux. Ainsi, la perception musicale, dans le contexte musical contemporain, ne se distingue guère de la perception des objets sonores et leurs articulations à plusieurs niveaux multi-échelles. Aux compositeurs reste la tache de faire sortir les objets sonores « du laboratoire » et de les permettre d'*émerger* dans ses *multiples projections isomorphiques* proportionnées par les œuvres musicales.

Dans nos futures recherches, nous prétendons donc développer des modèles informatiques simulant un apprentissage perceptuel des configurations et catégorisations morphologiques des objets sonores. Ces catégorisations sont basées sur les classifications *typo-morphologiques* proposées par Pierre Schaeffer [22]. Car, à notre avis, seulement une approche dans ce sens pourra s'occuper de la perception auditive de façon plus large et générale, en tenant compte pas seulement de la musique tonale mais aussi de la musique contemporaine, de la perception du langage et de la perception des scènes auditives environnantes. Ces essais seront, bien évidemment, interpolés avec nos futures investigations philosophiques sur la perception auditive et musicale.

6. REFERENCES

[1] Bayle, F. *Musique acousmatique – propositions... ...positions.* INA/E_d. Buchet/Chastel, Paris, 1993.

[2] Besnier, B. « La conceptualisation husserlienne du temps en 1913 », *Annales de phénoménologie*, Association pour la promotion de la phénoménologie, Beauvais, 2004.

[3] Bregman, A. *Auditory scene analysis : The perceptual organization of sound.* The MIT Press, Cambridge, MA, 1990.

[4] Chauviré, C. & Sackur, J. *Le vocabulaire de Wittgenstein.* Ellipses, Paris, 2003.

[5] Eco, U. *La structure absente.* Mercure de France, Paris,1972.

[6] Guillaume, P. *La psychologie de la forme.* Flammarion, Paris, 1979.

[7] Husserl, E. *Idées directrices pour une phénoménologie.* Gallimard, Paris, et Tel n° 94, 1950

[8] Husserl, E. *Leçons pour une phénoménologie de la conscience intime du temps.* PUF, Paris, 1964,

[9] Husserl, E. *Recherches Logiques*, vol. II, PUF, Paris, 1974.

[10] Husserl, E. *Sur la phénoménologie de la conscience intime du temps.* Jérôme Million, Grenoble, 2003.

[11] Husserl, E. « Psychologie et phénoménologie », *Annales de phénoménologie*, Association pour la promotion de la phénoménologie, Beauvais, 2004.

[12] Köhler, W. *Psychologie de la forme.* Ed. Gallimard, Paris, 2000.

[13] Kononen, T. *Self-organizing maps*. Springer-Verlag, Berlin, 1995.

[14] Kuhn, T. S. *La structures des révolutions scientifiques.* Flammarion, Paris, 1983.

[15] Leman, M. & Carreras, F. « Schema and Gestalt : Testing the Hyothesis of Psychoneural Isomorphism by Computer Simulaiton », *Music, Gestalt, and Computing*, Springer-Verlag, Berlin, 1997.

[16] Marion, J.-L. *Réduction et donation – Recherches sur Husserl, Heidegger et la phénoménologie.* PUF, Paris, 1989.

[17] Mattusch, U., « Emulating Gestalt Machanisms by Combining Symbolic and Subsymbolique Information Processing Procedures », *Music, Gestalt, and Computing*, Springer-Verlag, Berlin, 1997.

[18] Petitot, J. *Les catastrophes de la parole.* Maloine S.A., Paris, 1985a.

[19] Petitot, J. « Perception, cognition et morphologie », *La musique et les sciences cognitives*, Pierre Mardaga, Bruxelles, 1988.

[20] Petitot, J. *Physique du sens.* CNRS Editions, Paris, 2000.

[21] Petitot, J. « Neurogéométrie et phénoménologie de la perception », *Philosophies de la perception*, Odile Jacob, Paris, 2003.

[22] Schaeffer, P. *Traité des objets musicaux.* Seuil, Paris, 1966.

[23] Shepard, R. & Chipman, S. « Second-Order Isomorphism of Internal Representations : Shapes of States », *Cognitive Psychology n° 1*, 1970.

[24] Thom, R. *Stabilité structurelle et morphogénèse.* W. A. Benjamin, INC, Massachustts, 1972.

[25] Thom, R. *Paraboles et catastrophes.* Flammarion, Paris, 1983.

[26] Vaggione, H. « Son, temps, objet, syntaxe », *Musique, rationalité, langage – L'harmonie : du monde au matériau* in *Cahiers de philosophie du langage n°3*, L'Harmattan, Paris, 1998.

[27] Van Gelger, T. « Le fer en bois ? », *Naturaliser la phénoménologie.* CNRS Editions, Paris, 2000.

[28] Varela, F., Thompson, E., & Rosch, E. *Inscription corporelle de l'esprit*, Seuil, Paris, 1993.

[29] Varela, F. « Le présent spécieux », *Naturaliser la phénoménologie*, CNRS Editions, Paris, 2000.

[30] Wittgenstein, L. *Remarques sur la philosophie de la psychologie - Vol. I* . Editions T.E.R., Mauvezin, 1989.

[31] Wittgenstein, L. *Tractatus logico-philosophicus.* Gallimard, Paris, 1993.

[32] Wittgenstein, L. *Recherches Philosophiques.* Ed. Gallimard, Paris, 2004.