



HAL
open science

Dé-linéariser le son dans son rapport à l'image : émergence d'un nouvel objet audiovisuel complexe

Hervé Zénouda

► **To cite this version:**

Hervé Zénouda. Dé-linéariser le son dans son rapport à l'image : émergence d'un nouvel objet audiovisuel complexe. Caroline Angé. Les objets hypertextuels : pratiques et usages hypermédiatiques, Iste editions, pp.121-146, 2015, Systèmes d'information, web et société, 978-1-78406-096-1. sic_01759793

HAL Id: sic_01759793

https://archivesic.ccsd.cnrs.fr/sic_01759793

Submitted on 5 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Chapitre 4 : Dé-linéariser le son dans son rapport à l'image : émergence d'un nouvel objet audiovisuel complexe

4.1 Introduction

L'apport du son et de l'image fixe et animée a fait progressivement évoluer le modèle hypertexte (navigation réticulaire dans un ensemble de grains d'informations) vers celui d'hypermédia introduisant de problématiques nouvelles, en particulier, les questions liées à la sensorialité, aux sémiotiques multimodales et à la gestion de temporalités différentes (celle de la navigation dirigée par l'utilisateur et celles spécifiques de chacun des objets temporels) apparaissent. Le sonore, objet temporel par nature linéaire, se déployant dans l'espace sans représentation visuelle rend difficile sa manipulation directe par l'utilisateur d'un environnement interactif. Agir sur le son instaure dans le cadre des hypermédia, dispositif technique à but essentiellement informationnel ou communicationnel, une dimension instrumentale nouvelle où le geste de l'utilisateur est central. Celui-ci peut se déployer seul, enregistré par le biais de capteurs, de « Joy Stick » ou par le mouvement de la souris, ou dans une interaction avec (une action sur) l'image¹ à l'écran (image qui par sa dimension spatiale facilite une action directe sur celle-ci). La question d'une interaction avec le son nous amène alors vers des problématiques plus spécifiquement liées aux relations image/son, qui se structurent, dans une situation interactive, autour du tryptique image/son/geste. Le sens émergera alors des multiples relations entre ces trois pôles.

Nous avons montré, dans un précédent travail², que l'on pouvait repérer, au travers des différentes pratiques artistiques des siècles passés mettant en œuvre des relations entre l'image et le son, une dynamique de convergence des deux représentations. Cette dynamique qui s'inscrit dans la filiation de la synesthésie évolue historiquement de logiques de *correspondances* vers des logiques de *fusion* qui dans le numérique, vont se réaliser dans des structures de langage de plus en plus

¹ Voir « L'action sur l'image » (Harmattan, xxx)

² Mon livre Harmattan (Harmattan, 2008)

sophistiquées (langages de descriptions, méta-données...). Cette logique de fusion de l'image et du son dans ces structures de langages laissant entrevoir l'émergence d'un nouvel objet audiovisuel complexe en interaction avec son environnement qui dépasse les notions classiques d'image et de son. Il s'agira ici de comprendre en quoi cette non-linéarité du modèle hypertextuel et hypermédiatique de l'objet sonore, témoigne d'une construction spécifique du sens en amont dans la création par l'analyse d'œuvres mettant en lien des combinatoires plurielles et une expérience communicationnelle nouvelle. Ce chapitre dédié au rôle et à la place du sonore dans des dispositifs interactifs propose de retracer ce cheminement et d'explorer ce champ particulier des relations images/sons en s'appuyant sur plus de dix années de recherches et de pratiques personnelles dans le domaine du design sonore interactif.

4.2 Champ d'études

Les hypermédia présentés dans ce chapitre se trouvent au croisement de plusieurs approches et filiations : ouverture de l'œuvre, représentation du sonore telles que les partitions graphiques, interfaces graphiques de manipulation du son dans le contexte de l'informatique musicale, la sémantique des relations images/sons par les approches du design sonore dans le cinéma, la vidéo, le design sonore interactif et, les rapports entre composition et théorie des jeux, etc.

Nous pouvons distinguer trois grandes classes de relations image/son qui peuvent se côtoyer dans une même application ou bien structurer des classes d'applications spécifiques :

- *le son au service de l'image* (classe issue, pour le dire rapidement, de la pratique de l'audiovisuel classique et de l'illustration sonore),
- *l'image au service du son* (issue de la pratique de l'informatique musicale où l'image cherche à représenter du son),
- *l'image et le son dans une relation « équilibrée »* (situation plus spécifique aux créations hypermédiées).

La première classe, l'illustration sonore, est encore aujourd'hui la plus courante même si elle est enrichie par quelques pratiques interactives plus spécifiques au

support numérique. La seconde classe concerne des applications qui ont souvent la musique comme contenu (des systèmes d'apprentissage ou de documentation) ou des jeux à fortes composantes sonores (tables de mixage en tout genre etc). La relation est inversée, le sonore est l'élément premier et l'image un moyen de faciliter son accès. Les hypermédias apportent à ces applications une ergonomie et une interface graphique de plus en plus riche. La troisième, enfin, concerne les créations hypermédias. Celles-ci explorent et repoussent les limites techniques et esthétiques du numérique pour tenter de trouver un véritable langage spécifique. Les trois pôles de création, image, son et programmation y sont à égalité et sont travaillés en collaboration étroite dès la phase initiale du projet. Ces trois modes de relations entre les images et les sons peuvent, néanmoins, se côtoyer, de manière plus ou moins étroite, dans une même réalisation hypermédia.

Ces trois classes recouvrent un certain nombre de fonctions et d'applications que l'on peut catégoriser comme suit :

- *le son au service de l'image* : fond sonore, musique d'accompagnement, illustration sonore et bruitages, « feedback sonore » (essai/erreur), sonorisation de texte (fonction informative), sonification d'interface,
- *l'image au service du son* : navigation dans des paysages sonores, tables de mixage, jouets sonores³, nouveaux instruments⁴,
- *rapport image/son équilibré* : jouets sonores, nouveaux instruments, individuation d'un personnage interactif, correspondances de rythmes et expériences de synesthésies, monde ou être virtuel.

³ Un jouet sonore est un dispositif qui met en œuvre une interface graphique permettant de manipuler certains paramètres du sonore et du musical. Le dispositif permet ainsi un certain espace de jeu. À la différence des instruments de musique qui nécessitent un certain apprentissage, ces règles de jeu sont suffisamment fortes pour que le résultat sonore garde toujours une certaine cohérence.

La contrepartie négative de cette facilité de prise en main est à trouver dans la réduction importante des possibilités sonores de ce genre de dispositifs.

⁴ Un instrument est un dispositif produisant des événements sonores primaires. Un apprentissage minimum est nécessaire pour produire un énoncé musical cohérent. À la différence du jouet sonore, la contrepartie positive de cette difficulté d'apprentissage est dans la variété infinie des productions musicales produites.

4.3 Ancrages historiques

4.3.1 Ouverture de l'œuvre

La deuxième partie du vingtième siècle voit apparaître une volonté généralisée de libéralisation des formes artistiques. En littérature, le groupe OULIPO introduit l'utilisation intensive des mathématiques et des jeux sur les combinatoires dans l'acte d'écriture, tandis que des écrivains comme Joyce, Char, Michaux font exploser le récit et repoussent, à leurs extrêmes, les limites du langage. En 1965, Umberto Eco publie "L'œuvre ouverte" où il analyse l'ouverture comme le champ des interprétations possibles d'une œuvre artistique classiquement "fermée" dans son rendu mais aussi comme l'apparition de nouvelles formes d'œuvres où la notion d'ouverture est inscrite dans leur conception même. Il cite entre autre des œuvres musicales de compositeurs comme Stockhausen, Boulez, Pousseur ou Boucourechliev. Ceux-ci, en effet, sont les premiers à la fin des années 50 à proposer différentes approches pour envisager une composition musicale non plus comme une œuvre finie, figée, fixe mais comme un ensemble de possibles offrant à l'interprète un choix de parcours individualisés au moment de l'exécution.

Ce degré de liberté et d'autonomie que donne le compositeur à l'interprète mérite d'être précisé :

Liberté de décision pour un ou plusieurs paramètres sonores (intensité, durée, hauteur) soit totale soit par tirage aléatoire (ou autre) dans un ensemble de possibles prévu par le compositeur ("Sequenza I" de Berio).
Liberté de choix du nombre des parties (voies) indépendantes jouées (John Cage).
Liberté de modifier la forme de l'œuvre : soit en ayant recours à des parcours personnalisés dans l'œuvre soit en modifiant l'ordre même des séquences musicales ("Klavierstück XI" de Stockhausen, "Troisième sonate" de Pierre Boulez).
Liberté de choix du matériau (John Cage) : application de règles sur des matériaux sonores (ou autres) choisis par les interprètes.

Cette "ouverture" (mobilité) de la musique peut intervenir à différents niveaux et concerner les différents acteurs de l'acte musical :

Elle peut concerner le compositeur (qui se servira, dans ce cas, de cette mobilité pour pouvoir créer à partir d'une "œuvre processus" autant d'"œuvres objets" qu'il le souhaitera). Cette mobilité reste inaccessible à l'interprète qui jouera une des versions finalisées que lui procure le compositeur.

Elle peut concerner l'interprète (grâce à différentes techniques que l'on précisera dans la suite de ce chapitre). Cette mobilité reste cependant opaque à l'auditeur qui écoute à chaque audition une version de l'œuvre "fermée".

Elle peut concerner, en final, l'auditeur qui pourra, via un jeu d'interactions avec l'interprète, par son comportement (déplacement ou autre) modifier le cours de la musique soit au minima être conscient de la marge de liberté de jeu de l'interprète dans la composition (soit par une mise en scène "théâtrale" de cette marge de liberté soit que celle-ci soit délibérément audible dans la musique).
--

4.3.2 Nouveaux rapports images/sons

4.3.2.1 La musique audiovisuelle ou « colorée »

L'expression « musique audiovisuelle » est avancée dès le début des années 1940⁵ par John Whitney⁶ qui en définit les caractéristiques en travaillant la question de tempo et celle des rapports cinématographiques spécifiques (couples temps/espace et forme/couleur). Néanmoins, la « musique audiovisuelle » ou « colorée » est présente bien plus tôt, dès les années 1920, avec des réalisateurs comme Oskar Fischinger⁷ ou Len Lye⁸ et s'inscrit dans la mouvance des interrogations sur la correspondance entre les arts et sur la synesthésie chère à des peintres abstraits comme Kandinsky ou Paul Klee. La volonté de ces réalisateurs est de trouver un langage

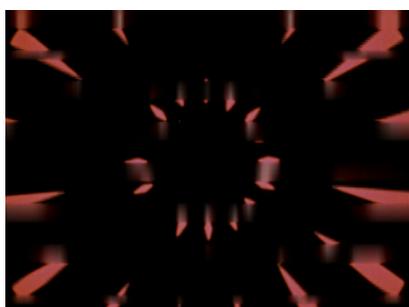
⁵ *Five film exercises* (1943 - 1944) (<http://www.my-os.net/blog/index.php?2006/06/20/330-john-whitney>)

⁶ Whitney John (1980), *Digital harmony on the complementary of music and visual arts*, Bytes Books, New Hampshire.

⁷ « *Studie Nr 7. Poème visuel* » (1929-1930), « *Studie Nr 8* » (1931) ...

⁸ « *A Colour Box* » (1935) ...

cinématographique spécifique « libéré » de la narration et des formes issues du roman et du théâtre. Pour cela, ces cinéastes vont s'éloigner de la fonction première de captation du réel du cinéma en choisissant d'utiliser de préférence des formes non figuratives et en recherchant des modèles de construction de l'œuvre du côté de cet art abstrait par excellence qu'est la musique. Avec la durée et le rythme comme dénominateur commun, ces réalisateurs vont chercher à produire des œuvres synesthésiques « totales » jouant sur les relations subtiles entre les différents sens.



*Figure 1 : Jordan Belson (1961), Allures
(<http://www.mediafire.com/?fy920bhvu6q6b1v>)*

4.3.2.2 Les partitions graphiques

Les principes d'indétermination développés par de nombreux compositeurs dès la fin des années cinquante⁹ obligent à repenser la partition comme outil de communication musicale. Un genre nouveau apparaît dans un rapprochement des arts sonores et visuels. Les compositeurs élargissent leurs palettes de symboles (couleurs, formes, profondeurs, textures...) pour décrire de nouvelles sonorités et des processus musicaux inédits. Ces nouvelles représentations introduisent des espaces de liberté liés à l'improvisation et mettent en avant des indications musicales globales ou volontairement floues qui sapent l'obsession de contrôle de la musique contemporaine. Le compositeur se charge de la macro structure (tandis que l'interprète se charge du détail de la réalisation), et communique à l'interprète d'autres types d'informations moins précises comme des sensations, des émotions, des mouvements, des dynamiques.

⁹ John Cage mais aussi Earle Brown, Pierre Boulez, André Boucourechliev et bien d'autres....

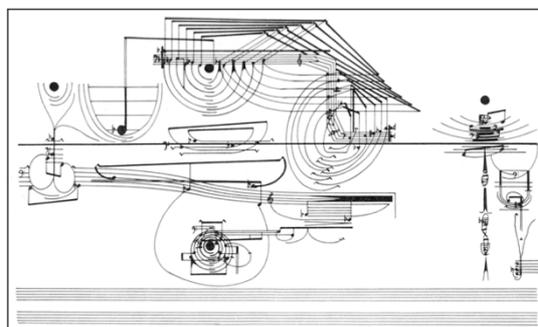


Figure 2 : Cornelius Cardew (1963 - 1967), *Treatise*

4.3.2.3 Les interfaces graphiques pour la composition musicale

Dans les années soixante-dix, le compositeur Iannis Xenakis, intéressé lui aussi par les rapports entre la vision et l'audition, propose une interface graphique pour la composition musicale sous le nom de l'UPIC^{10 11}. Le compositeur peut, grâce à une palette graphique, dessiner les formes d'ondes et les enveloppes d'amplitude, et contrôler à la fois la structure du son et la forme globale de l'œuvre. Plus récemment, Golan Levin¹², chercheur au MIT¹³, propose une série d'« instruments »¹⁴ qui associent de manière très étroite l'image, le son et le geste. Développeur de palettes graphiques qu'il sonorise en générant en temps réel des sons de synthèse, il associe des paramètres du geste, comme la direction et la vitesse du mouvement ou la pression du crayon électronique, à des paramètres sonores, comme le timbre, la hauteur, le panoramique, et à des paramètres graphiques, comme la couleur, l'épaisseur de trait ou la direction. Chez Golan Levin, les notions d'interactivité et de générativité sont étroitement liées : l'image et le son sont générés en temps réel, à partir du geste de l'utilisateur créant ainsi une « substance

¹⁰ Unité Polyagogique Informatique du CEMAMu (Centre d'Etudes de Mathématiques et Automatique Musicales)

¹¹ <http://www.youtube.com/watch?v=yztoaNakKok>

¹² <http://acg.media.mit.edu/people/golan/>

¹³ Massachusetts Institute of Technology (USA)

¹⁴ « *Aurora* » (1999), « *Floo* » (1999), « *Yellowtail* » (1999), « *Loom* » (1999), « *Warbo* » (2000)

audiovisuelle malléable générative et interactive »¹⁵. En France, la société Blue Yeti¹⁶ propose le système « *Grapholine* » de dessin musical à doubles écrans basé sur la transformation d'échantillons sonores (via une banque de sons standards mais aussi personnalisables) par la synthèse granulaire et offrant une large gamme de manipulations et de relations image/son (vitesse du tracé, transparence, luminosité, couleur, pression du crayon...). À l'UTC de Compiègne, deux étudiants¹⁷ en dernière année de cursus informatique, proposent une application de réalité virtuelle (*Immersive Music Painter, 2 010*¹⁸) où l'interactant dessine dans l'espace, par son geste, des courbes de différentes couleurs et d'épaisseurs auxquelles est associé un son ou une mélodie dont la hauteur, le panoramique et le volume sont contrôlables.

4.3.3 Théorie des jeux et composition musicale

Si on trouve une influence de la théorie des jeux dans la composition chez de nombreux compositeurs (Mozart¹⁹...), celle-ci est devenue plus présente au fur et à mesure que l'industrie des jeux vidéo a pris son envol et trouve son marché dans une série d'audio-games accessible sur I-phone et tablette à très bas prix. Le terme de « jouabilité »²⁰ est utilisé pour décrire l'ensemble des éléments de l'expérience vidéo-ludique : à la fois règles, interface et maniabilité du jeu ainsi que leurs appropriations par le joueur (le plaisir retiré et sa facilité d'appropriation). Si l'on définit le terme de jouabilité comme le caractère de ce qui est jouable, de ce qui peut être joué, alors on peut trouver une correspondance musicale dans la notion de musicalité, de ce qui est « bien écrit » pour un instrument donné et qui « sonne » bien. On pourra alors parler de la « jouabilité » d'une pièce pour piano en référence

¹⁵ « *an inexhaustible audiovisual substance which is created and manipulated through gestural mark-making* » Golan Levin, *Painterly Interfaces for Audiovisual Performance, B.S. Art and Design*, [LEVIN 1994], p.19.

¹⁶ <http://www.blueyeti.fr/Grapholine.html>

¹⁷ Camille Barot et Kevin Carpentier.

¹⁸ <http://www.utc.fr/imp/>

¹⁹ « Jeux musicaux, K294d », Mozart, 1800.

²⁰ Le terme de « game-play » est aussi utilisé dans le monde anglo-saxon.

au plaisir du pianiste à jouer cette partition et à son aptitude à investir son interprétation. Concernant la dimension de la prise en main des règles du jeu, un compositeur comme John Cage a remarquablement déplacé les affectations de ces rôles et permis un rapprochement significatif vers la dimension ludique. En effet, dès la fin des années cinquante, il enrichit sa recherche sur le hasard et l'indétermination en transformant la partition en un espace de jeu mettant en œuvre des règles que l'interprète doit incorporer et surtout faire vivre. La partition passe ainsi du statut de *notation de paramètres musicaux* à celui de *notations d'actions*. Pour exemple, dans « *Variations I* » (1958), « *Variations II* » (1961) ou « *Cartridge Music* » (1960) : les notes sont inscrites en valeurs relatives (grosseur de points se référant à l'intensité et à la durée) sur des feuilles transparentes que l'interprète superpose à d'autres feuilles elles aussi transparentes contenant des règles de jeux (des formes géométriques, un cadran d'horloge...). Cette superposition de feuilles contenant des informations de niveaux différents (notes et règles) produit la partition à interpréter. Nous le voyons ici, l'interprète doit, avant de jouer la moindre note, agencer un dispositif pour produire les informations musicales à jouer, un « travail » classiquement alloué au compositeur lui-même. Cette nouvelle définition du rôle de l'instrumentiste permet de retrouver les dimensions des deux termes anglais qui décrivent le jeu : « game » et « play ». Le terme « game » soulignant la notion de structure (système de règles que le joueur s'impose de respecter pour mener à bien son action), tandis que le terme « play » soulignerait l'attitude ludique adopté par le joueur (« l'action menée par celui qui joue ») (Genvo, 2002). En modifiant le rôle de l'interprète, John Cage augmente considérablement l'espace de liberté et de créativité de celui-ci, deux dimensions à la base de l'attitude ludique.

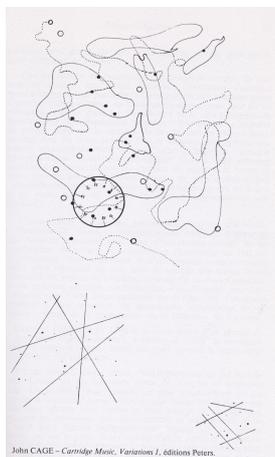


Figure 3 : John Cage (1960), Cartridge Music

Ce rapprochement de la composition et de la théorie des jeux se retrouve de manière encore plus explicite chez le compositeur dans sa célèbre pièce/performance « Reunion » (1968) qui le met en scène avec Marcel Duchamp dans une partie d'échecs où chaque mouvement de pièce dirige de manière plus ou moins aléatoire des événements musicaux.



Figure 4 : John Cage (1968), Reunion
(avec Marcel Duchamp à Toronto)

Nous retrouvons la théorie du jeu chez d'autres compositeurs comme Iannis Xenakis « Duel » (1959), « Stratégie » (1962) et « Linaia Agon » (1972), Mauricio Kagel (« Match » (1964) inspiré d'un match de tennis) ou John Zorn « Cobra » (1984).

4.4 Analyse d'œuvres

4.4.1 Créations visuelles et sonores

Legato, *Cellos* et *Moon tribe*²¹ sont trois réalisations qui mettent en scène des personnages filaires de danseurs. À partir de ces mêmes éléments graphiques, chacune de ces trois réalisations exploite un aspect musical bien particulier : *Legato* met en place un mixage de lignes mélodiques sur un thème en boucle, *Cellos* ordonne différentes mélodies préétablies, *Moon tribe* permet de synchroniser des boucles rythmiques. Ces trois propositions de Nicolas Clauss s'inscrivent dans la catégorie des jouets sonores²². En effet, la structure même de l'interaction est calquée ou transposée de structures musicales comme l'harmonie, le contrepoint, l'ordonnement de mélodie, la synchronisation de rythmes. L'image n'est pas ici une représentation du sonore, mais possède sa propre cohérence esthétique installant ainsi des effets de correspondance arbitraires avec le son. De même, chacun des modes, visuel et sonore, possède son propre temps ; le temps de l'animation graphique et le temps musical ne se recouvrant pas toujours parfaitement produit ainsi des décalages des boucles temporelles visuelles et sonores. Le geste ne fusionne pas les deux modes mais les coordonne. Il est le point de rencontre, la frontière entre ces deux mondes. Si cette relation provoque une fusion sensorielle, elle ne peut être que fugitive, non-univoque et soumise aux interprétations variant assez fortement d'un utilisateur à l'autre. Elle n'est pas directement inscrite dans le processus technique mais plutôt réalisée, au détour d'un geste intempestif, comme un plus d'appropriation de l'utilisateur.

²¹ <http://www.flyingpuppet.com>.

²² Un jouet sonore est un dispositif qui met en œuvre une interface graphique permettant de manipuler certains paramètres du sonore et du musical. Le dispositif permet ainsi un certain espace de jeu. À la différence des instruments de musique qui nécessitent un certain apprentissage, ces règles de jeu sont suffisamment fortes pour que le résultat sonore garde toujours une certaine cohérence.



Figure 5 : « Moon Tribe »

La série interactive appelée *Zoo*²³ propose un bestiaire interactif où chaque animal traité possède des propriétés interactives particulières. Chacun des tableaux propose quelques principes communs simples : un animal traverse l'écran, l'interactant peut alors le saisir, le manipuler, le disloquer, bref découvrir ses propriétés cachées. La manipulation du personnage déclenche des modifications sonores et visuelles qui peuvent modifier radicalement l'atmosphère initiale. Le son participe à l'identité du personnage au même titre que ses autres attributs (visuels ou interactifs). L'évènement sonore déclenché va ainsi jouer sur la manière dont l'utilisateur manipulera le personnage qui elle-même sera influencée par les modifications visuelles et sonores ainsi générées. Cet évènement sonore peut être décrit comme une mélodie interne liée au personnage interactif et est constitutive du dialogue qui se noue avec l'utilisateur puisque directement modifié par ce jeu. Si dans les tableaux de Nicolas Clauss, on ordonnait et mixait des éléments sonores entre eux via l'image sans pouvoir modifier leurs intégrités, ici l'interactivité modifiera directement la matière audiovisuelle. Des modifications des paramètres sonores vont être ainsi liées aux modifications visuelles induites par le geste (modification du volume, de la vitesse de lecture et de hauteur de notes, de la position spatiale).

²³ Frédéric Durieu (programmation, image), Jean Jacques Birgé (musique), www.lecielestbleu.com, 2001-2002.

Le CD-rom dédié à Léopold Senghor²⁴ propose un système de musique générative qui permet des variations infinies pour un même écran. Le sommaire est sonorisé par une musique dont le mixage est modifié par le déplacement du curseur et le survol des icônes dédiés aux différents chapitres. Une base de données sonore permet d'associer de nombreux sons à une même interaction²⁵.

Développeur de palettes graphiques qu'il sonorise en générant en temps réel des sons de synthèse, Golan Levin²⁶, associe des paramètres du geste comme la direction et la vitesse du mouvement ou la pression du crayon électronique à des paramètres sonores comme le timbre, la hauteur, le panoramique et, graphiques comme la couleur, l'épaisseur de trait ou la direction. Chez Golan Levin, les notions d'interactivité et de générativité sont étroitement liées : l'image comme le son sont générés en temps réel, à partir du geste de l'utilisateur. Avec *Yellowtrail*, le comportement de l'animation dépend strictement de la forme dessinée et de la vitesse du mouvement de l'utilisateur. Levin utilise ensuite ces graphismes comme spectrogramme inversé²⁷ pour générer le son. Un spectrogramme est la représentation graphique d'un son défini par ses informations de temps et d'amplitude. La technique de spectrogramme inversé permet de faire le chemin inverse, partir d'une image pour générer un son²⁸. Dans *Loom*, les paramètres du geste sont appliqués aux paramètres de la synthèse en modulation de fréquence²⁹. Ainsi, la vitesse du trait contrôle l'amplitude, la pression du crayon influe sur le vibrato du son ainsi que son amplitude, les courbes du trait influent sur la modulation. Chaque objet graphique animé a sa propre tête de lecture et contient donc son propre temps permettant de créer de riches effets de polyrythmie. Avec *Aurora*, l'interactant dessine des nuages animés de couleur qui obéissent au geste.

²⁴ Senghor, Jériko, Paris 1999.

²⁵ *La musique du cédérom Léopold Sédar-Senghor*, <http://perso.orange.fr/roland.cahen/>

²⁶ <http://acg.media.mit.edu/people/golan/>

²⁷ « Pattern playback ».

²⁸ Voir des logiciels du type de *Metasynt*.

²⁹ Popularisée par le DX7 dans les années 1980.

Ces nuages sont constitués de centaines de filaments de couleur qui sont associés aux nombreux grains sonores générés par une synthèse de type granulaire. Si dans *Loom*, le contrôle de la synthèse en modulation de fréquence nécessitait quelques dizaines de paramètres, dans *Aurora*, la synthèse granulaire en nécessite plusieurs milliers ! Il s'agit donc de faire correspondre les données des nuages visuels à la multitude de données que nécessite cette synthèse. De plus, cette mise en relation doit être signifiante au niveau perceptif pour l'utilisateur et l'on sait que distribuer directement les données du domaine visuel au domaine du sonore, ne donne pas souvent le résultat escompté. Pour créer un espace de traduction entre ces deux ensembles de données, Levin fait appel à des techniques de distribution statistique pour regrouper, augmenter ou réduire le nombre d'informations et les associer de manière rationnelle et signifiante.

4.4.2 Processus et Expériences communicationnelles

Antoine Schmitt³⁰ est un artiste numérique qui propose des œuvres centrées sur une communication avec des entités dotées de comportements. L'activité de programmation au cœur de son travail met en avant la notion de processus au détriment d'un travail plus traditionnel sur les images et les sons. Ainsi, Antoine Schmitt va utiliser des moyens visuels et sonores volontairement réduits pour mieux concentrer ses œuvres sur les causes de ces émergences médiatiques : le comportement de l'entité avec laquelle nous entrons en communication. Chaque œuvre est donc tout d'abord à comprendre comme une expérience communicationnelle entre un interactant et une entité douée de comportements et exprimant sa « présence » par le biais d'images et de sons générés entièrement par le programme. Cette perception du comportement est principalement rendue par le mouvement et l'interrogation sur cette cause peut se résumer à « pourquoi ça

³⁰ <http://www.gratin.org/as/>.

bouge ? ». La phase de conception/réalisation, elle, est abordée comme une approche de simulation qui cherche à expérimenter des lois physiques imaginaires³¹.

4.4.3 Audio-games

« *Aura* » (sur iPhone et iPad)³², très proche de l'esthétique de pièces comme « *Allegretto* » de Fishinger (1936) ou « *Allures* » de Belson (1961), propose, sur un fond musical généré, de construire ses propres mélodies (toujours en harmonie avec le fond sonore) en choisissant la note, le timbre et le volume via des formes colorées simples à l'écran. Le processus de création audio-visuel se fait de manière fluide et ininterrompue et le fond musical et visuel change en secouant simplement son iPhone. Le joueur produit, en touchant l'écran de son doigt, musiques et formes abstraites dans ce qui est présenté comme un générateur interactif de musique (et d'image) d'ambiance. Une application comme « *SynthPond* »³³ (iPhone et iPad) propose de positionner des formes sur des cercles de diamètres différents. Des notes sont jouées quand ces formes rencontrent les ondes visuelles produites par des formes carrées ou l'intersection des autres cercles. À la différence d'« *Aura* », « *SynthPond* » produit une musique de boucles répétitives dont il est possible de prévoir visuellement le rendu sonore en suivant le déplacement des formes et leur rapprochement des différentes intersections. Le joueur peut sélectionner les hauteurs, le timbre et le tempo des boucles permettant ainsi un résultat sonore potentiellement complexe.

³¹ *Le travail du temps : programmer « un mode d'être, un entretien avec Antoine Schmitt »*, Samuel Bianchini, <http://www.gratin.org/as/txts/index.html>.

³² <http://www.youtube.com/watch?v=rb-9AWP9RXw&feature=related>

³³ http://www.youtube.com/watch?v=mN4Rig_A8lc&feature=related

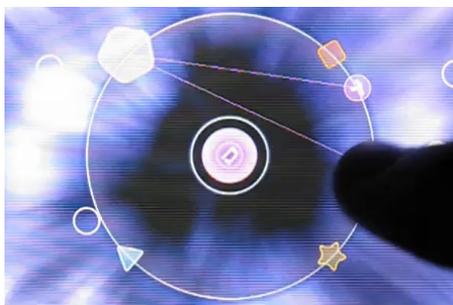


Figure 6 : « Aura »



Figure 7 : « Synthpond »

Deux jeux originaux de tirs musicaux : « Rez » (Mizuguchi, 2001, Sega)³⁴ et « Child of Eden » (Mizuguchi, 2010, Xbox360)³⁵ nous rapprochent de jeux plus classiques en gardant une dimension sonore importante. « Rez » se présente comme un jeu de combat où la partie musicale est créée dynamiquement en fonction des actions du joueur. Chaque tir du joueur ou des vaisseaux ennemis produit un son qui s'adapte rythmiquement et harmoniquement à la bande son. L'expérience de jeu devient ainsi particulièrement originale, la mission ludique (détruire des vaisseaux ennemis) est la source de la production de l'image et du son. Le discours du concepteur faisant référence à Kandinsky et à la synesthésie relie ce jeu aux jeux décrits précédemment. Pour accentuer cette recherche de synesthésie, l'agglomérat de sensations simultanées est stimulé par la vibration de la manette de jeu au

³⁴ <http://www.youtube.com/watch?v=2a1qsp9hXMw>

³⁵ http://www.youtube.com/watch?v=xuYWLYjOa_0&feature=fvst

moment de la synchronisation de l'image et du son. Dans « *Child of Eden* », Tetsuya Mizuguchi continue cette quête d'une jouabilité générant le son en utilisant les innovations technologiques les plus récentes (image Haute Définition, son spatialisé en 5.1, reconnaissance de mouvements avec Kinect permettant de ne plus utiliser ni souris ni manette). La reconnaissance des mouvements permet une prise directe avec les images ainsi que de nouveaux types d'interactions (le claquement de mains, par exemple, permet de changer d'armes pendant le jeu). L'esthétique visuelle y est beaucoup plus chatoyante et psychédélique que dans *Rez* et la mission ludique est de supprimer un virus qui cherche à mettre en péril le projet de recréer une personnalité humaine au sein d'Eden (mémoire de l'humanité). Chaque niveau correspond à un palier de cette archive de l'humanité et le dernier niveau intègre les contributions personnelles des joueurs sous la forme de photos de leurs moments heureux.



Figure 8 : « Rez »



Figure 9 : « Child of Eden »

« *Metris* »³⁶ (Mark Havryliv, Terumi Narushima, Applet Java) est une adaptation du célèbre jeu « *Tetris* » à une dimension de création musicale. À chaque block en déplacement est associé un son de cloche généré par de la synthèse en temps réel. Les actions de l'utilisateur qui déplace ou qui applique des rotations sur ce block produisent des modifications sonores de hauteurs. La manière dont ces blocks s'encastrent produit différents accords. L'ensemble des possibilités permet de produire une composition micro-tonale sophistiquée sans perdre l'intérêt du jeu initial de « *Tetris* ».

« *Pasy02* »³⁷ (iPhone et iPad) se présente comme une grille que l'on peut déformer et étirer à loisir et qui revient à la normale dans un effet d'élasticité. Ces modifications physiques influent sur les paramètres comme la hauteur, le tempo et le timbre de boucles musicales jouées par un synthétiseur dont les formes d'ondes (sinusoïde, triangle, carré, dents de scie) peuvent être sélectionnées par le joueur pour obtenir des résultats sonores variés. Le dispositif très simple offre une approche particulièrement originale et directe de production du son.

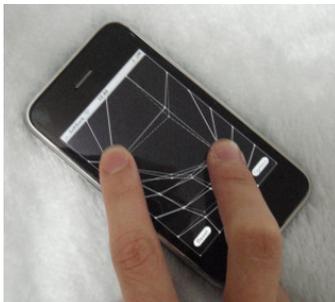


Figure 10 : « *Pasy02* »

Enfin, « *Elektro-Plankton* »³⁸ (Toshio Iwai, Nintendo) est un jeu emblématique qui propose un ensemble de dix tableaux mettant en scène des planctons dans des

³⁶ Havryliv Mark, Narushima Terumi, « *Metris: a game environment for music performance* », <http://ro.uow.edu.au/era/313/>

³⁷ <http://www.youtube.com/watch?v=JmqdvxLpj6g&feature=related>

³⁸ <http://www.youtube.com/watch?v=aPkPGcANA1g>

situations musicales variées³⁹ agrémentées de graphismes ludiques. Certains tableaux soulignent le lien étroit entre geste musical et geste pictural. Ainsi avec « *Tracys* », le joueur trace des lignes (droites ou courbes) que le plancton suivra en jouant des notes de piano au rythme de la création de ligne. D'autres mettent en scène la dimension labyrinthique et rhizomique de la musique (parcours dans un espace de notes possibles) : avec « *Luminaria* », plusieurs planctons se déplacent sur une grille faite de noeuds et de liens, chaque lien étant orienté. Le joueur peut changer les connexions entre les noeuds et modifier ainsi les trajectoires des planctons, transformant par là même les notes émises. D'autres mettent en avant un dispositif proche de ce que nous pourrions appeler une installation sonore : les « *Hanenbrows* », par exemple, sont projetés dans l'écran et rebondissent sur des feuilles qui produisent une note musicale. En inclinant ces feuilles, le joueur produit des rebonds différents et influe sur les mélodies générées. La fréquence de sollicitation des feuilles modifie leurs couleurs et leurs tonalités. Un tableau comme « *Nanocarps* », met en scène un ensemble de planctons qui ont leurs comportements et déplacements propres, le son est produit par leurs entrechoquements et l'utilisation du micro permet au joueur de réorganiser leurs mouvements. De même, le tableau « *Volvoice* » utilise le micro de l'ordinateur pour capturer des sons dont les paramètres (hauteur, vitesse, filtre) peuvent ensuite être modifiés à loisir par la déformation à l'écran du plancton associé. Enfin, le tableau « *Sun-Animalcule* » propose de déposer des œufs sur l'écran. De la position choisie dépendra la note associée. La lumière d'un cycle jour/nuit fait éclore les œufs qui libèrent leur note associée puis les embryons planctons vont grandir et faire évoluer leur comportement musical.

³⁹ Dix tableaux : Tracy/Trapy, Hanenbow, Luminaria, Sun-Animalcule, Rec-Rec, Nanocarp, Lumiloop, Marine-Snow, Beatnes, Volvoice.

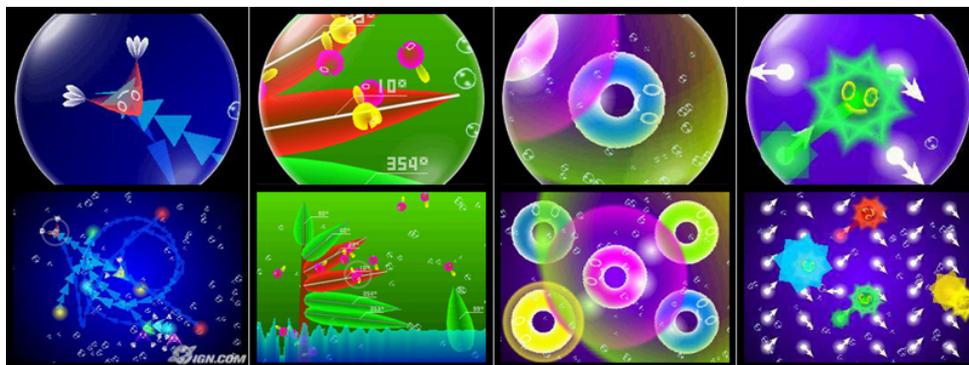


Figure 11 : « *ElectroPlankton* »

4.5 De nouveaux effets perceptifs

Ces différents exemples nous permettent d'appréhender de nouveaux effets perceptifs spécifiques au média numérique. Les hypermédia ancrent le son à l'image de manière plus forte qu'au cinéma en accentuant les synchronisations intempestives dues à la manipulation interactive. Cet apparent asservissement du son à l'image ne doit pas masquer la transformation de cette même image au contact rapproché du son. Ainsi dans les œuvres qui proposent des jeux sonores et visuels riches, le rapport entre le son et l'image s'inverse souvent. En manipulant une image, on produit du son. Après quelques instants, l'intérêt sonore et l'intérêt visuel viennent s'entremêler de manière à ne plus savoir si c'est le son ou l'image qui dictent les mouvements du geste. Les effets de synchronisations se multiplient dans une synchrèse⁴⁰ généralisée liant images, sons et gestes. De plus, cette synchronisation s'effectue aussi dans le sens inverse, de la substance audiovisuelle vers l'interacteur. Dans le cas de la manipulation d'objets temporels comme l'image animée ou le son, l'interacteur manipule ces objets mais se synchronise aussi sur eux en adaptant son geste à leurs propres temporalités. Il s'agit ici d'un effet de *syntonie*⁴¹ qui a été

⁴⁰ Mot valise créé par Michel Chion (ref) associant les termes synchronisation et synthèse pour désigner l'effet de fusion perceptive créée par la synchronisation temporelle de l'image et du son.

⁴¹ Syntonie : « égalité de fréquence des oscillations libres (de deux ou plusieurs circuits) [...] circuits en syntonie : accordés sur la même longueur d'ondes. Syntonisation : [...] Réglage de résonance qui assure le rendement maximum. » (Le Petit Robert).

étudié, par exemple, par Jean Louis Weissberg⁴². D'autre part, en associant plusieurs sons possibles à une image, les hypermédia matérialisent l'axe paradigmatique développant ainsi un déroulement temporel particulier lié à l'interaction. Selon les règles d'association de ces sons (par calcul, tirés aléatoirement, par famille de sons), se développent différentes significations et une perception globale de l'ensemble de ces sons dans leurs rapports à l'image. Le sens se construit ainsi autant *entre* les sons et les images d'un même ensemble que dans leurs significations spécifiques. Enfin, dans les systèmes interactifs qui mettent en jeu des fonctions génératives ou des situations de communication avec des entités ou des mondes simulés, la cause des émergences médiatiques (les règles de production dans le premier cas, le modèle de simulation dans le second) incite l'interacteur à dépasser les effets perceptifs de ces images et sons générés pour percevoir le processus *en amont*. Un processus qui n'est pas abstrait mais qui s'incarne dans la substance audiovisuelle imbriquant ainsi les différents niveaux de perception. C'est la spécificité des hypermédia d'associer étroitement ces aspects du sensible et du calcul⁴³ dans une même expérience. On peut dire que, de même que le défilement linéaire des photogrammes au cinéma fait apparaître le mouvement, la manipulation interactive fait apparaître le processus.

4.6 Eléments pour une méthode d'analyse

Nous proposons dans cette partie des pistes pour élaborer une méthode d'analyse des hypermédia :

- Distinction des relations entre l'image et le son : trois modes de relations se côtoient et glissent de l'un à l'autre dans une même hypermédia : le son au service

⁴² Jean Louis Weissberg , *Corps à corps – à propos de la morsure*, <http://hypermedia.univ-paris8.fr/seminaires/semaction/seminaires/txt01-02/fs-02.htm>, Mars 2002.

⁴³ Lev Manovitch, dans son livre *The language of new media* (MIT Press, 2001) utilise les expressions de « *couche culturelle* » et de « *couche computationnelle* » pour désigner ces deux mondes. Pour lui, les nouveaux médias devenus programmables (« *cinématographique en surface, digital dans leur structure et computationnel dans leur logique* »), appellent ainsi une nouvelle approche qui va chercher ses modèles dans le domaine de l'informatique.

de l'image (issue de l'illustration sonore et de l'audiovisuel classique), l'image au service du son (issue de l'informatique musicale où l'image cherche à représenter le son et à le manipuler), l'image et le son dans une relation égale (plus spécifique aux « audio-games » et aux hypermédia) produisant des effets de fusion perceptibles entre les deux modalités.

- Distinction des modes de déclenchements des sons : avec des sons déclenchés par une intervention directe de l'utilisateur par déplacement de la souris, clic ou survol d'une icône ou manipulation d'image, des sons déclenchés automatiquement par le système sans intervention de l'utilisateur comme les fonds sonores automatiques et enfin des sons déclenchés automatiquement par le système mais liés à une action de l'utilisateur comme, par exemple, à un parcours particulier dans l'application ou à un temps passé sur une interaction. Ces différents modes de déclenchements ont tendance à se parasiter et à provoquer une certaine confusion. Ainsi examiner précisément comment l'auteur gère les mixages intempestifs de ces différentes couches sonores est une piste précieuse pour une analyse détaillée des rapports images/sons dans une situation interactive.

- Prise en compte du triptyque image/son/geste : qui s'exprime dans la notion de mapping (la transcription d'informations captées dans un registre, le mouvement ou la manipulation graphique, sur un autre registre comme ici le musical). Nous distinguons généralement plusieurs types de mapping : la relation où à un paramètre d'un domaine correspond un paramètre de l'autre (*one-to-one*), celle où un paramètre d'un domaine est associé à plusieurs paramètres de l'autre (*one-to-many*), et enfin la relation où, à l'inverse, plusieurs paramètres d'un domaine sont associés à un seul paramètre de l'autre (*many-to-one*). Dans ces associations multi-sensorielles, la question du choix des différents paramètres des modalités en jeu et la manière dont ils sont reliés entre eux est primordiale. Nous pouvons, en effet, repérer sur quelles dimensions sonores et visuelles l'interaction s'applique et quels effets perceptibles sont associés. Concernant le son : la note (hauteur, durée, intensité), le timbre (enveloppe, formes d'onde), les règles de génération du discours musical, le mixage de plusieurs pistes, les paramètres généraux comme le tempo et à quels

paramètres de l'image ceux-ci sont affectés (couleur, forme, opacité, netteté, cadre, niveau d'iconicité). Quels effets de sens sont produits par ces multiples possibilités de liaisons image/son/geste ?

- Analyse d'objectifs cognitifs différents : Les hypermédia mettent en place des situations particulières où le geste de l'utilisateur contrôle et produit dans le même temps de l'image et du son tout en participant à une logique supplémentaire, celle des objectifs de l'application elle-même (détruire des vaisseaux spatiaux pour « Rez » par exemple, ou agencer des obstacles pour gérer des rebonds de balles pour « *Elektroplankton* »). Nous avons vu que cette situation spécifique produit des effets perceptibles complexes où les effets de synchronisation et de fusion de l'image et du son sont démultipliés par le geste et des enjeux cognitifs différents. La grande difficulté sera de relier ces différents niveaux (la logique de la manipulation de l'image, règles du jeu, production musicale) de manière signifiante (ni dans une relation de conflit ni dans une relation trop distante ou anecdotique). L'idéal étant que les règles de manipulations se transforment en nouveau geste instrumental dans la production du sonore et soit ainsi susceptibles de produire de la nouveauté musicale. Pour obtenir un résultat optimal, les règles et les missions de l'application doivent donc éclairer les structures musicales en jeu tout en gardant leur impact et leur cohérence en tant que jeu. On voit bien là, toute la difficulté de cet équilibre souhaité.

- Ces nouveaux objets soulignent la nécessité de développer des approches de sémiotiques multi-modales pour les analyser, approches qui prennent en compte à la fois le visible, l'audible et le geste.

Quelques outils peuvent nous aider à avancer dans ce sens :

- En 1961, Abraham Moles⁴⁴ propose une *échelle d'iconicité* en treize échelons pour classer les images en fonction de leur degré de ressemblance avec l'objet réel représenté. Cet axe progressif part des représentations les plus

⁴⁴ Moles A. (1972), « *Théorie de l'information et perception esthétique* », Denoel-Gonthier, Paris.

concrètes et figuratives pour aller vers les représentations les plus éloignées ayant un degré d'abstraction élevé comme les langages naturels ou langages artificiels (mathématiques etc.). Cette échelle d'iconicité peut être appliquée au son en se développant sur deux axes : celui qui va du concret au musical et celui qui va du capté au simulé (synthèse sonore).

- Inspiré par la sémiotique de Peirce pour l'image, le compositeur François Bayle⁴⁵ définit trois propriétés du son liées à l'intention d'écoute : l'*icône* (image *isomorphe* ou *im-son*) : l'objet est dénoté par l'ensemble de ses caractères, l'*indice* (image *indicielle* ou *di-son*) : certains traits de l'image dénotent l'objet, le *symbole* (une *métaphore* ou *mé-son*) : l'image représente l'objet par associativité. Ces trois natures de signes débouchent sur trois niveaux d'écoute : celui où les sons s'écoutent comme étant rattachés à des référents directement identifiables de la réalité (*qualité* : *quali-signe*) ; celui où la relation est plus abstraite, le son devient un élément significatif de quelque chose. Écoute spécialisée : Les sons s'écoutent comme transformés (filtrage, transposition, insert...), indices d'un travail du compositeur (*singularité* : *syn-signe*) ; et enfin, celui où le signe répond à une loi connue qui est en dehors du signe lui-même. (rebond, oscillation...). Écoute tournée vers le sens de l'organisation, de la loi formelle... (*stabilité* : *légi-signe*).

- Conçues au groupe de recherche musicale de Marseille, l'*Unité Sémiotique Temporelle* est « un segment musical qui, même hors contexte, possède une signification temporelle précise due à son organisation morphologique ».⁴⁶ Dix-neuf UST ont été ainsi répertoriées et nommées par des étiquettes : un mot ou une expression littéraire qui décrit le plus directement possible la manière dont se déploie l'énergie sonore dans le temps⁴⁷, le plus souvent à l'aide d'une « *appellation*

⁴⁵ Bayle F. (1993), *Musique acousmatique : propositions ... positions*, Buchet Castel, Paris.

⁴⁶ MIM, *Les UST : éléments nouveaux d'analyse musicale*, [MIM 2002], p°18.

⁴⁷ Chute, Contracté, Etendue, Elan, En flottement, En suspension, Etirement, Freinage, Lourdeur, Obsessionnel, Par vagues, Qui avance, Qui tourne, Qui veut démarrer, Sans direction par divergence d'information, Sans direction par excès d'information, Stationnaire, Sur l'erre, Suspension-interrogation, Trajectoire inexorable.

métaphorique, un qualificatif qui fait référence à quelque chose d'extra musical »⁴⁸. Cette référence extra-musicale est un premier pont vers une généralisation de ces étiquettes à d'autres modalités. Dans ce sens, on peut souligner que l'ensemble de ces étiquettes fait référence à un mouvement énergétique ou spatial qui les relie naturellement au geste et à l'image.

4.7 Proposition d'une typologie

Enfin, nous pouvons proposer une typologie la plus simple et la plus générale possible afin d'intégrer tous types d'œuvres hypermédia à dimensions musicales ainsi que tous types de musiques produites indépendamment des styles et du niveau d'expertise des utilisateurs. Pour cela, nous nous concentrons sur la question de la représentation (fixe ou dynamique) de deux types d'entités sonores : les éléments de base (axe paradigmatique) et les éléments construits par la manipulation (la jouabilité) proposée par le jeu (axe syntagmatique).

- L'axe paradigmatique vertical concerne la représentation des éléments sonores de base que l'objet met à disposition du joueur. La représentation graphique de cet élément de base pourra s'étendre du solfège classique à des créations graphiques sans lien direct avec le sonore en passant par des représentations abstraites liées à la synesthésie (et à tous types de correspondances avec les différents paramètres du son). Notons que la frontière entre un son élémentaire et un son faisant l'objet d'une construction dépend du type de jeu et de l'interaction proposée. Ainsi, ce qui sera un objet sonore composé dans un hypermédia peut être pris comme élément de base dans un autre. Les items de cet axe pourront donc être modifiés en fonction des objets à analyser et seront positionnés suivant leurs éloignements ou proximités du solfège traditionnel.

- L'axe syntagmatique horizontal concerne la représentation d'objets sonores construits par la jouabilité

⁴⁸ *Idem* p. 36.

du jeu. Il s'agit bien ici de représentations des opérations de constructions de ces objets sonores de second niveau. Ces représentations peuvent être fixes et proches des manipulations musicales classiques (représentations des opérations de répétitions, de transpositions, de renversements...) ou dynamiques (représentations de simulations de phénomènes physiques comme le rebond, l'explosion, l'accumulations...) ou encore décrire des dynamiques de jouabilités « musicologiques » ou « a-musicologiques ». Les premières peuvent trouver des analogies musicales directes (déplacements dans l'espace, jeux de miroirs...), les secondes n'ont aucune analogie avec le musical et impliquent donc des relations arbitraires et ad-hoc liées à la diégèse du jeu (jeux de combats...). Comme pour l'axe paradigmatique, les items de l'axe syntagmatique pourront être modifiés en fonction des oeuvres à analyser et seront positionnés suivant leurs éloignements ou proximités du solfège traditionnel.

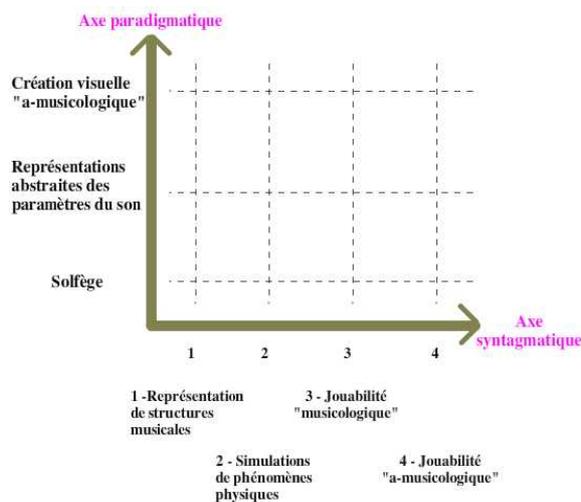


Figure 5 : Typologie d'œuvres hypermédia à dimensions musicales

4.8 L'objet complexe multimodale

Notre hypothèse principale présentée dans notre introduction avance que la technologie numérique, par de nombreuses caractéristiques (codage numérique

commun, représentations communes, interactivité, synchronisation de l'image et du son par le geste), fusionne l'image et le son en un nouvel objet audiovisuel que l'on peut qualifier de complexe et qui repose sur le programme sous-jacent. Chacun de ses aspects comme l'image, le son, ses comportements devient alors une facette particulière de cet objet sensible que l'on ne peut plus qualifier d'image (qu'elle soit interactive, actée⁴⁹ ou autre⁵⁰) ni de son.

Cette fusion progressive des deux modalités se fait, dans le numérique, dans des structures de langage de plus en plus élaborées : langages de description et de production, métalangages, métadonnées, langages a-média. Cette primauté renvoie à la profonde filiation entre langage et informatique. Celle-ci se révèle, en effet, inhérente au paradigme informatique lui-même, paradigme plus large que la simple question de la numérisation. Si la numérisation, c'est-à-dire la transformation d'objets hétérogènes analogiques en fichiers de même nature : une suite de 0 et de 1, représente bien une étape primordiale de normalisation entre les objets (une véritable *transduction* dans le sens qu'en donne Simondon⁵¹), elle permet surtout à ceux-ci d'être, à la fois, manipulables par du calcul et d'inclure en leur sein des structures réflexives de description. Ces calculs peuvent ainsi s'appliquer à chaque strate de l'objet numérique : son codage binaire et ses couches de descriptions langagières plus ou moins évoluées.

Ainsi l'environnement numérique ne se contente pas d'ajouter de nouvelles fonctionnalités à des images et des sons qui garderaient leur intégrité et identité préalables en en faisant des images ou des sons augmentés par du calcul et des comportements. Au contraire, cet environnement favorise l'émergence d'une

⁴⁹ Jean-Louis Weissberg dans « *Présences à distance* » [WEISSBERG 1999] utilise l'expression « *image actée* » pour souligner la dimension participative de l'interactant, la trace du geste dans la communication avec l'image.

⁵⁰ Sophie Lavaud dans son article « *Les images/systèmes : des alter ego autonomes* » [LAVAUD 2006] utilise l'expression « *image système* » pour souligner la dimension complexe des nouvelles images interactives et autonomes. L'image devient la porte d'entrée d'un ensemble d'interactions, de manipulations et de transformations complexes qui intègrent une certaine dose d'imprévisibilité.

⁵¹ Gilbert Simondon développe, dans son livre *Du mode d'existence des objets techniques*, [SIMONDON 1989], le concept de transduction : passage et mutation d'un objet dans des environnements divers.

nouvelle forme médiatique qui dépasse les notions classiques d'image et de son et qui nécessite de nouveaux outils d'analyse pour être étudiée. Le point de rupture avec l'image et le son entendus dans leur acception traditionnelle nous paraît se trouver dans un renversement de perspective dû à la reconnaissance de ce nouvel objet comme processus. Il ne peut, de ce fait, être considéré comme une simple image ou un simple son, mais plutôt comme un complexe d'images et de sons en émergence régi par des règles de production, d'association et de transformation.

Bien sûr, cet objet est en final perçu par un utilisateur. Et en tant qu'objet de perception, il n'existe qu'au travers de sensations visuelles et sonores. Néanmoins, de nombreuses relations *entre* les images, *entre* les sons ainsi qu'*entre* ces ensembles d'images et de sons forcent l'utilisateur à dépasser les manifestations médiatiques temporaires pour se construire des repères à un niveau supérieur, un niveau méta.⁵² De même, la précarité⁵³ de ces images et de ces sons, due à leur mobilité et transformations perpétuelles, réduit la force de chaque image et de chaque son et renforce le poids du système logique qui les produit et les relie. Cette nouvelle situation perceptive ne va pas sans une expérience particulière qui renvoie à la perception du processus mis en œuvre. Ainsi les aspects dynamiques et réflexifs de ce nouvel objet nous paraissent plus caractéristiques que les manifestations médiatiques résultantes du processus. Celles-ci peuvent être considérées comme des couches sensibles du programme qui appellent l'utilisateur à remonter vers le code :

« Le numérique articule ainsi des univers hétérogènes de représentations où ce qui est perçu, vu, entendu, senti n'est qu'une des formes temporaires d'expression d'un imperceptible, celui des flots construits de données numériques homogènes qui sans cesse le manifestent

Tout objet de surface du numérique, qu'il se révèle dans l'univers des images, du texte, du son ou d'un mélange de ces univers, est en lui-même

⁵² Sur le métarécit, voir l'article de Pierre Barboza : « Sale temps pour la fiction » in, Desgouttes J.P. (Sous la direction de), *Mise en scène du discours audiovisuel*, [BARBOZA 1999].

⁵³ Pierre Barboza, « L'extrême précarité de l'image numérique » in *L'image actée : scénarisations numériques*, [BARBOZA 2006].

dynamique c'est-à-dire formé d'un ensemble fluctuant, et dont les rapports de fluctuations sont l'objet de calculs constants, de codes, de signes, de temps, de relations. [...] Parce que l'art numérique démontre que ce qui n'est pas perceptible existe au moins au même titre que ce qui l'est, le percepteur de la manifestation d'art numérique est en mesure de percevoir à la fois l'au-delà et l'en-deçà de sa perception ».⁵⁴

Nous remarquons que ce code, cet imperceptible, se fait de plus en plus sensible par la logique et la dynamique des relations en jeu. Une expérience médiatique propre au numérique est donc à trouver dans une communication avec un objet complexe qui devient sensible au travers un ensemble d'images et de sons ou de matrices génératives d'images et de sons.

Ainsi nous pouvons avancer une description de notre nouvel objet audiovisuel complexe et structuré comme un objet autonome qui contient des données, des comportements, des capacités de productions médiatiques (sonores, visuelles, textuelles) ; des comportements interactifs qui gèrent la relation avec l'utilisateur, une capacité de mémorisation, une capacité de modifier dans le temps l'ensemble de ces composants (données, comportements, représentations etc.). De plus, il peut avoir un but ou une intention autonome. Dès lors, c'est un *processus multimodal en interaction avec son environnement*, à la fois numérique et humain, qui est constitué d'un enchevêtrement de plusieurs niveaux et de plusieurs dimensions.

Quatre grandes caractéristiques sont à retenir, ce nouvel objet est en effet :

- *autonome* car il contient ses propres données et ses propres programmes,
- *complexe* parce qu'hétérogène et parce que l'ensemble est plus que la somme de ses composants,
- *multimodal* parce qu'il « *articule des univers hétérogènes de représentations* »⁵⁵,

⁵⁴ Jean Pierre Balpe, *Les contextes de l'art numérique*, [BALPE 2000], p.105 et p.139.

⁵⁵ *Idem*, p.105

- *structuré* en métadonnées qui sont organisées sur plusieurs niveaux enchevêtrés.

Le terme *complexe* renvoie à l'approche système pour analyser ce processus en émergence. *Objet autonome* renvoie à la programmation par objets et par agents comme implémentation informatique de l'approche système. Le terme *structuré* renvoie aux approches privilégiant les métadonnées dans la description et la production. Le terme *multimodal* renvoie au statut fondamentalement multi-code et multi-représentation de cet objet.

Le nouvel objet audiovisuel se structure en de nombreuses couches qui vont se superposer les unes aux autres dans un continuum de structures formelles de plus en plus proches des logiques humaines. Partant du codage binaire, elles évoluent vers des représentations cognitives de haut niveau. De plus, ces structures de données réduisent de manière très sensible les frontières entre langages de description et langages de production⁵⁶ ainsi qu'entre sons et images calculés d'un côté et sons et images captés de l'autre. D'autre part, l'*autonomie* du processus nous semble être une notion primordiale pour percevoir cet objet dans son intégrité. En effet, si la perception ne se limite plus aux émergences médiatiques résultantes, elle doit trouver ses repères face à une entité qui a une cohérence, une intention, un but et qui se présente comme un objet autonome contenant ses propres données et ses propres méthodes d'action.

Par ailleurs, une des caractéristiques de notre nouvel objet audiovisuel est aussi à trouver dans son double statut d'objet génératif *et* interactif : à la fois capable d'une autonomie lui permettant de générer ses propres représentations visuelles et sonores, et dans le même temps tourné vers un dialogue avec un utilisateur. Les règles de génération sont alors soumises aux manipulations de l'interacteur soit dans des relations simples quand la réaction est alors directement associable à l'action, soit dans des relations plus complexes, c'est-à-dire décorrélées, augmentées ou diminuées.

⁵⁶ Avec les avancées du traitement du signal et des langages de synthèse d'images et de sons orientés contenus.

4.9 Synthèse conclusive

Nous avons cherché, dans cet article, à souligner l'impact de l'introduction du son et de l'image fixe et animée au modèle hypertexte dans son passage aux hypermédia. Après avoir présenté plusieurs œuvres, nous avons cherché à les classer, à en montrer les interactions images/sons spécifiques ainsi que les nouveaux effets perceptifs qui en résultent. Nous avons ensuite proposé des éléments pour une mise en œuvre d'une grille d'analyse de ces nouveaux objets. Nous avons enfin souligné que l'interactivité rapproche l'image et le son jusqu'à le fusionner dans le geste utilisateur permettant ainsi la possible émergence d'un nouvel objet audiovisuel pour lequel, le processus devient central et sensible. Ce nouvel objet audiovisuel complexe nous paraît indiquer le dépassement des notions classiques d'image et de son tel que nous les avons connu jusqu'à présent.

4.7 Bibliographie

- [BAL 00] BALPE J.P., Contextes de l'art numérique, Editions Hermes, Paris, 2000.
- [BON ROU 03] BONARDI A., ROUSSEAU F., « "Music-ripping" : des pratiques qui provoquent la musicologie », dans *ICHIM 03*, 2003.
- [BOS 98] BOSSEUR J-Y., Musique et arts graphiques, Interactions au XXe siècle, Minerve, Paris, 1998.
- [CAG 76] CAGE J., Pour les oiseaux, Belfond, Paris, 1976.
- [GEN 02] GENVO S., « Le game design de jeux vidéo : quels types de narration ? », dans *Transmédiabilité de la narration vidéoludique : quels outils d'analyse ?*, Comparaison, Peter Lang, 2002, p. 103-112, 2002.
- [DEL 01] DELALANDE F., « La seconde révolution technologique de la musique occidentale », dans *Le son des musiques (entre technologie et esthétique)*, Buchet/Chastel, 2001.
- [FLU 06] FLUSSER V., Pour une philosophie de la photographie, Circé, Paris, 1996.
- [KAN 69] KANDINSKY V., Du spirituel dans l'art, et dans la peinture en

particulier, Denoël-Gonthier, Paris, 1969.

- [MAN 01] MANOVITCH L., *The language of new media*, MIT Press, Cambridge, 2001.

- [STR 01] STRANSKA L., *Les partitions graphiques dans le contexte de la création musicale Tchèque et Slovaque de la seconde moitié du vingtième siècle*, Thèse de Musicologie, Paris IV, 2001.

- [POP 85] POPPER F., *Art, Action et participation*, Paris, Klincksieck, 1985.

- [VIN 99] VINET H., *Concepts d'interfaces graphiques pour la production musicale et sonore in Interfaces homme-machine et création musicale*, Hermes, Paris, 1999.

- [VIN 04] VINET H., « Explorer la musique à l'âge numérique », dans *Les cahiers de médiologie N° 18 Révolutions industrielles de la musique*, Fayard, 2004.

- [ZEN 08] ZENOUDA H., *Les images et les sons dans les hypermédias artistiques contemporains : de la correspondance à la fusion*, Paris, Editions L'Harmattan, 2008.