



**HAL**  
open science

# Que fait le numérique à la relation entre image et son dans les arts contemporains ?

Benoît Montigné

► **To cite this version:**

Benoît Montigné. Que fait le numérique à la relation entre image et son dans les arts contemporains ?. Art et histoire de l'art. Université Panthéon-Sorbonne - Paris I, 2023. Français. NNT : 2023PA01H300 . tel-04411283

**HAL Id: tel-04411283**

**<https://theses.hal.science/tel-04411283>**

Submitted on 23 Jan 2024

**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.



## THÈSE

Pour obtenir le grade de  
**Docteur de l'Université Paris 1 Panthéon-Sorbonne**  
**Discipline : Arts plastiques, Esthétique et Sciences de l'Art**

Soutenance le 4 juillet 2023  
par  
**Benoît MONTIGNÉ**

# Que fait le numérique à la relation entre image et son dans les arts contemporains ?

Sous la direction de Monsieur le Professeur Jean-Marie Dallet,  
Professeur des universités à l'Université Paris 1 Panthéon-Sorbonne

### **MEMBRES DU JURY**

**Monsieur Patrick NARDIN**, Professeur des universités, Université Paris 8 Vincennes Saint-Denis, pré-rapporteur

**Monsieur Jean-Paul FOURMENTRAUX**, Professeur des universités, Aix-Marseille Université, pré-rapporteur

**Madame Karine LE BAIL**, Historienne et chercheuse, CNRS/EHESS, examinatrice

**Madame Giusy PISANO**, Professeure des universités, École nationale supérieure Louis-Lumière, examinatrice

**Monsieur Frédéric SOJCHER**, Professeur des universités, Université Paris 1 Panthéon-Sorbonne, examinateur

**Monsieur Jean-Marie DALLET** Professeur des universités, Université Paris 1 Panthéon-Sorbonne, directeur

## Résumé

Avec l'avènement de ce que l'on a appelé de manière diffuse les nouveaux médias — qui incluent les arts dits numériques — ont émergé dans l'art contemporain de nouvelles formes hybrides, *audio-visuelles*, animées par une présence plus ou moins forte du numérique. Dans ces formes, le sonore, souvent oublié dans les études théoriques, est devenu — tout comme l'image — génératif et interactif. Lorsque nous évoquons le sonore, nous ne parlons pas uniquement de musique, mais bel et bien d'un matériau aussi porteur de sens que l'image ; plus encore, faisant sens avec l'image. Nous avons donc réuni et étudié des œuvres, à caractère numérique, dans lesquelles le son et l'image dialoguent dans un rapport d'équité. Notre champ d'étude se focalise sur trois territoires artistiques qui lient les débuts du cinéma sonore aux formes numériques contemporaines : la *musique visuelle* programmée, les nouvelles dramaturgies interactives qui s'appuient sur des bases de données, et un ensemble de travaux artistiques mettant en jeu le visage et la voix. En émergent des processus renouvelés ou inédits, de la sonification au *Deep Learning*.

Ce travail d'analyse a mis en lumière différentes figures *audiographiques* élémentaires, qui ont constitué le point de départ de notre recherche artistique. Nous avons pour cela élaboré une plate-forme d'expérimentation en ligne, qui présente des séries d'études interactives liant image animée abstraite et son. Celles-ci visent à repartir des fondements de chacun de ces deux champs, à les croiser, afin d'en dégager une grammaire plastique transversale.

À travers notre analyse et notre travail de recherche-crédation, c'est le périmètre du dialogue contemporain entre image et son à l'ère du numérique que nous souhaitons contribuer à circonscrire. L'idée étant d'en extraire ce qui pourrait relever de modalités fondamentalement nouvelles, aux côtés de relations déjà établies dans l'histoire des arts.

## Mots-clés

Interaction audio-visuelle — Art numérique — Art sonore — Musique visuelle générative — Art audiographique — Cinéma et bases de données — Intelligence artificielle — Visualisation et sonification — Agents conversationnels — Surveillance — Reconnaissance vocale et faciale — Perception multimodale

## **Abstract**

With the advent of what we loosely call digital media, which includes the so-called Digital Arts, we see emerging in contemporary art new hybrid forms, *audio-visual*, animated by a more or less digital presence. Through those forms, sound, often forgotten in theoretical studies, has become — just like the image — generative and interactive. When we evoke sound, we are not only talking about music but an element that carries just as much sense as the image, even more so, making sense of the image. We have collected and studied digital work in which sound and image work together equitably. Our field of studies focalizes on three artistic territories that link the beginnings of sound films to contemporary digital forms: programmed *Visual Music*, new interactive dramaturgies arts which rely on a database, and a compilation of artistic works using the face and the voice. Out of this emerge new or innovative processes, from *sonification* to *Deep Learning*.

This analytical work brings to light different elemental *audiographic* figures, that constitute the starting point of our artistic research. For this, we have elaborated an online experimental platform that presents a series of interactive studies linking animated abstract images and sound. The objective is to establish a foundation for plastic transversal grammar between the two fields.

Throughout our analyses and research-creation work and during the digital era, we wish to contribute to circumscribing the perimeter for contemporary dialogue between image and sound. The idea is to extract what could raise new fundamental modalities next to already established relations in Art history.

## **Keywords**

Audiovisual interaction — Digital Art — Sound Art — Generative Visual Music — Audiographic Art — Database Cinema — Artificial intelligence — Visualization and sonification — Chatbots — Surveillance — Voice and facial recognition — Multimodal perception

## Remerciements

Le travail de recherche est un long chemin. Je souhaite avant tout remercier celui qui en a été le guide : mon directeur de thèse, Jean-Marie Dallet. En contrepoint d'échanges toujours passionnants, il a remarquablement suivi mon travail et accompagné sa rédaction. Il m'a également permis de contribuer à l'organisation d'événements scientifiques. Sa science de l'interactivité, associée à ses qualités humaines, m'a été d'une aide considérable pour mettre en œuvre la construction de mes propres *architectures de mémoire*.

Je remercie également l'École doctorale APESA de l'Université Paris 1 Panthéon-Sorbonne et son équipe, tout particulièrement sa directrice Sandrine Morsillo, ainsi que son administratrice Shahida Minot.

Plusieurs artistes, chercheurs ou commissaires ont très gentiment répondu à mes questions, merci à eux : Peter Beyls, Larry Cuba, Philippe Esling de l'Ircam, Kyriaki Goni, Cindy Keefer du Center for Visual Music, Barbara Lattanzi, Maria du studio Moritz Fehr, Jerry Reed, John Whitney Jr., Anne Zeitz.

J'adresse également des remerciements à Roberto Barbanti, de l'Université Paris 8 Vincennes-Saint-Denis, qui m'a permis de lancer cette recherche, à Emmanuel Mahé pour m'avoir intégré à l'Ensadlab, à Serge Verny et à tous les membres du groupe de recherche *Formes du mouvement* pour nos discussions si souvent captivantes.

J'adresse des remerciements particuliers à Gilles Alvarez, directeur de la Biennale Némó, grâce à qui j'ai pu développer et présenter ma recherche-crédation.

Je remercie également Françoise Courbis de son soutien permanent.

D'autres personnes m'ont permis de présenter mon travail lors de colloques ou de journées d'études, ou bien d'échanger avec eux. Je les en remercie : Laurence Allard de la Sorbonne Nouvelle, David Christoffel, Odile Fayet du Centre Pompidou, Gaétan Nael du festival Maintenant, Daniel Pressnitzer de l'ENS. Merci également à Matthieu Saladin de m'avoir confié des charges de cours en arts audiovisuels à Paris 8.

Une pensée aussi pour Alain Bioteau, disparu, et les souvenirs de notre cours en binôme, autour des musiques expérimentales et des arts sonores à la Sorbonne Nouvelle.

Merci à mes parents, qui m'ont toujours laissé suivre mes chemins expérimentaux.

Enfin, et probablement est-ce là le plus important, je tenais à remercier ma femme Véronique, pour ses encouragements, son infinie patience, et ses relectures précieuses. Ce travail n'existerait pas sans elle. Nos deux enfants, Constance et Victorien, ont eux aussi, à leur manière, soutenu la réalisation de ce travail ; je les en remercie.

## Sommaire

<b>Introduction.....</b>	<b>8</b>
--------------------------	----------

### **Première partie.**

#### **Formes de la musique visuelle programmée : de la composition d'abstractions**

<b>visuelles et sonores à l'info-esthétique.....</b>	<b>19</b>
--	-----------

1.1 Abstractions graphiques et musicales du début du XXe siècle.....	21
1.2 Le son des premiers films programmés.....	43
1.3 Inventer ses outils : des interfaces artistiques pour un art audiovisuel.....	71
1.4 Visualisation et sonification des données.....	100

### **Deuxième partie.**

#### **Dramaturgies interactives, dialectique des images et des sons dans les média**

<b>numériques.....</b>	<b>124</b>
------------------------	------------

2.1 Programmer l'axe-temps des images-son : inverser, ralentir, accélérer.....	126
2.2 Épuisement du fragment audiovisuel : du saut interactif à la microstructure.....	145
2.3 Esthétique des bases de données d'images et de sons.....	164
2.4 Les images et les sons des paysages technologiques.....	199

### **Troisième partie.**

#### **Dialogues et choralités. Les visages-voix de l'ère numérique contemporaine.....** **217** |

3.1 Du portrait interactif à l'agent conversationnel.....	219
3.2 Visages et voix diagrammatiques.....	241
3.3 Choralités machiniques.....	262
3.4 Talking Heads.....	279

## **Quatrième partie.**

<b>Carnet de recherche-création. Une plate-forme expérimentale web pour la recherche audiographique .....</b>	<b>296</b>
4.1 Des études pour l'exploration de l'art « audiographique ».....	297
4.2 Un objet informatique audiographique : approche conceptuelle et technique .....	304
4.3 Quelques exemples d'études.....	310
4.4 Des interfaces : façonnage, navigation et composition.....	326
4.5 Une version dédiée au jeune public : musigraphicode.....	333
<b>Conclusion.....</b>	<b>339</b>
<b>Bibliographie.....</b>	<b>348</b>
<b>Sitographie.....</b>	<b>356</b>
<b>Index nominum.....</b>	<b>363</b>
<b>Index rerum.....</b>	<b>368</b>
<b>Table des illustrations.....</b>	<b>373</b>
<b>Extraits et documentations vidéo.....</b>	<b>382</b>
<b>Table des matières.....</b>	<b>387</b>
<b>Annexes.....</b>	<b>391</b>

## ***Site web compagnon***

Nous avons donné, pour la majorité des œuvres constituant le corpus de cette thèse, des liens vers des extraits vidéo. Les adresses web (URL) sont à la fois présentes en note de bas de page, et réunies sur une page dédiée en fin de document. Beaucoup de copies illégales de films sont accessibles sur le Web, mais, afin de ne pas en favoriser la circulation, nous avons tâché de privilégier des captures publiées sur les sites des artistes eux-mêmes, ou bien d'institutions culturelles.

Les extraits sélectionnés sont également publiés sur un mini-site compagnon, que nous avons réalisé, et qui est accessible à l'adresse suivante :

<http://quefaitlenerique.benoit-montigne.net/>

## Introduction

### « Terme manquant : **sonore** »

Avec l'avènement de ce que l'on a appelé de manière diffuse les nouveaux média — qui incluent les arts dits numériques —, ont émergé dans l'art contemporain de nouvelles formes hybrides, audio-visuelles, animées par une présence accrue, à différents niveaux, des technologies numériques. Dans ces formes, comme l'analyse très justement Jean-Paul Fourmentraux<sup>1</sup>, les images sont devenues algorithmiques, interactives, performatives, *responsive*<sup>2</sup> ou bien opératoires<sup>3</sup>. Les sons — entrelacs de bruits, de voix, de musiques —, quand ils habitent ces formes, n'en sont pas moins devenus, eux aussi, numériques, algorithmiques, interactifs, synthétiques et connectés. Pourtant, ils semblent être laissés de côté dans la majorité des analyses contemporaines dédiées au champ du numérique, au profit des nouvelles images.

Cette situation n'est pas nouvelle. Au début des années 1990, dans l'introduction de son essai *L'audiovision*, le théoricien Michel Chion rappelait combien les analyses du cinéma avaient, pour une majorité d'entre elles, éludé la question du son<sup>4</sup>. Le projet de son ouvrage consistait à reconsidérer les questions de l'image et du son au cinéma par le biais d'une méthodologie analytique audio-visuelle — c'est-à-dire plaçant l'image et le son dans un rapport d'équité — en dégagant des figures qui illustrent leurs influences réciproques du point de vue de celui qu'il nomme l'audio-spectateur. Il y analyse les influences du son sur la perception du mouvement ou sur celle du temps, définit différents types d'écoute, explore la synchronisation et la question du hors-champ. Malgré ce travail important (et de plusieurs autres qui ont suivi<sup>5</sup>), il subsiste toujours une prédominance de l'image sur le son.

1 Jean-Paul Fourmentraux (dir.), *Images interactives. Art contemporain, recherche et création numérique*, Bruxelles, La Lettre volée, coll. « Essais », 2017.

2 Le *responsive design* désigne une manière de travailler l'adaptativité des productions éditoriales et artistiques numériques en fonction de leur différents modes de visualisation des écrans (téléphone, tablette, ordinateur).

3 Le terme fait référence à la notion d'image opératoire développée par l'artiste allemand Harun Farocki. Ces images, déployées notamment dans le champ militaire ou industriel, sont celles de la surveillance et du contrôle qui s'exercent dans un contexte machinique, sans parfois même le regard de l'être humain.

4 Michel Chion, *L'audio-vision. Son et image au cinéma*, Paris, Armand Colin, 1990, p. 3.

5 Citons, parmi d'autres : Laurent Jullier, Claude Bailblé ou Philippe Langlois pour le son au cinéma.

Lors de nos explorations web, lorsque nous avons essayé de fabriquer des requêtes textuelles favorisant des rencontres entre les deux domaines de l'image et du son, les moteurs de recherche nous ont d'ailleurs fréquemment renvoyé la mention « Terme manquant : sonore ».

Le sonore s'est pourtant imposé au fil du XX<sup>e</sup> siècle comme un matériau de plus en plus important, par l'action mutuelle de deux territoires jusque-là antagonistes : d'un côté le domaine musical, de l'autre les arts plastiques.

La musique, tout d'abord, s'est vue confrontée à de multiples mutations. En 1913, Luigi Russolo signe son manifeste *L'art des bruits*, dans lequel il imagine une musique fondée sur les bruits, par « la jonction et la substitution des bruits aux sons<sup>6</sup> ». Le compositeur avant-gardiste Edgar Varèse, dans la cadre d'une théorisation de la musique au cinéma, va proposer, en place de l'expression « musique » (qui lui semble perdre sa signification) la notion plus large de « son organisé<sup>7</sup> ». Dans les années 1950, Pierre Schaeffer fait émerger en France la musique concrète<sup>8</sup> tandis que naît en Allemagne la musique électronique analogique. À partir des années 1960, depuis les *glissandi*<sup>9</sup> de Iannis Xenakis jusqu'à la musique spectrale<sup>10</sup> de Gérard Grisey, en passant par la recherche-composition du son de synthèse informatique explorée par Jean-Claude Risset<sup>11</sup>, la musique contemporaine va continuer d'étendre son territoire par l'exploration du sonore. D'autres exemples devraient être cités, mais ceux-ci témoignent déjà des mutations, nombreuses, qu'a pu connaître le champ musical.

Cette émergence du son dans la musique, que le musicologue Makis Solomos a remarquablement documenté dans son ouvrage *De la musique au son*<sup>12</sup>, est venue se confronter parallèlement à une autre émergence, celle du sonore dans l'art. Dans les années 1960, le matériau sonore est devenu central chez des artistes tel que Max

6 Luigi Russolo, *L'art des bruits. Manifeste futuriste 1913*, Paris, Allia, 2003, p. 25.

7 Edgar Varèse, *Écrits*, Paris, Christian Bourgois, 1983, p. 106.

8 Une musique fondée sur l'« écoute réduite », c'est-à-dire une écoute du son pour lui-même, déconnecté de sa source, considéré comme un objet sonore.

9 Des glissements, c'est-à-dire des descentes de notes de musique dans le grave ou des montées dans l'aigu qui semblent continues pour la perception, par le jeu instrumental. Produits collectivement dans un orchestre, ces ensembles de *glissandi* peuvent former des masses sonores.

10 Travail sonore sur le timbre et les propriétés acoustiques du son.

11 Jean-Claude Risset, *Écrits — Volume I. Composer le son : repères d'une exploration du monde sonore numérique*, Paris, Hermann, coll. « GREAM/Création contemporaine », 2014, p. 76.

12 Makis Solomos, *De la musique au son. L'émergence du son dans la musique des XX<sup>e</sup>-XXI<sup>e</sup> siècles*, Rennes, coll. « Aesthetica », Presses Universitaires de Rennes, 2013.

Neuhaus<sup>13</sup>, interrogeant ainsi l'existence d'un « art sonore ». En effet, comme le montrent les premières expositions qui propagent ce terme<sup>14</sup>, les frontières sont parfois floues entre ce que l'on pourrait nommer les musiques expérimentales (le compositeur John Cage pouvant en être la figure pionnière) et les arts du son. Nous comprenons la position d'artistes ou de théoriciens qui évitent l'utilisation des termes « art sonore » (*Sound Art* en anglais) ou « arts sonores ». Le théoricien et historien du sonore Douglas Kahn préfère par exemple parler de « son dans l'art<sup>15</sup> ». Nous pensons toutefois que ces termes — quand bien même les artistes eux-mêmes ne souhaitent pas s'en réclamer — permettent de remettre en lumière l'importance de ce matériau. Le nombre croissant d'ouvrages consacrés au sonore ces dix dernières années en montrent le bénéfice. De la même manière, l'exploration de ce champ par l'archéologie des média<sup>16</sup> a également fait émerger les *Sound Studies*<sup>17</sup>, qui nous aident à mieux comprendre les conditions culturelles d'émergence des média sonores.

Ce rééquilibrage nécessaire entre le sonore et l'image par le développement des arts et des études peut toutefois provoquer, comme pour le visuel, une forme de focalisation cloisonnée sur le champ du sonore. C'est pourquoi nous ne cherchons pas à dire si nous explorons le son dans les arts visuels ou le visuel dans les arts du son. Nous abordons l'histoire des formes audio-visuelles en nous positionnant à des nœuds, des croisements entre arts du son et arts de l'image, en donnant à chacune des composantes une place, un poids égal, pour en analyser les relations. Nous souhaitons contribuer à définir le périmètre de ce dialogue contemporain, dans la perspective d'en extraire ce qui pourrait relever de modalités fondamentalement nouvelles, aux côtés de relations déjà persistantes dans l'histoire des arts.

---

13 Max Neuhaus a été l'un des pionniers du travail artistique avec la matière sonore. Musicien, plasticien, il était l'inventeur du terme « installation sonore », mais ne souhaitait pas lui-même se revendiquer comme un artiste sonore.

14 L'exposition *Sonic Boom* à la Hayward Gallery en 2000 à Londres, dont le commissaire était David Toop, incluait les œuvres d'artistes plasticiens utilisant le son comme matériau, mais aussi les pièces de musiciens généralement associés aux musiques expérimentales. Pour différents théoriciens, ce type de programmation a contribué à un mélange des genres.

15 Douglas Kahn, « Sound art, Art, Musique », in *Tacet n° 3*, « De l'espace sonore », Dijon, Haute école des arts du Rhin/Les Presses du réel, coll. « OH CET ECHO », 2014, p. 332.

16 Nous écrivons le terme média sans le « s », pour éviter la confusion avec les médias d'information.

17 Terme générique regroupant les différentes approches pluridisciplinaires visant à circonscrire les cultures sonores.

## Les relations entre image et son dans l'art

Si l'approche audio-visuelle reste encore, selon nous, trop rare dans les écrits consacrés aux formes contemporaines des arts médiatiques<sup>18</sup>, il émerge tout de même des réflexions nourries de cette transversalité. Parmi les rares grandes expositions en France, citons *Sons & Lumières*<sup>19</sup> en 2004 au Centre Pompidou à Paris, sous-titrée « Une histoire du son dans l'art du XX<sup>e</sup> siècle ». Elle documente de nombreux points de rencontre entre les sons et les abstractions visuelles, depuis l'histoire des orgues à couleur<sup>20</sup> jusqu'aux expérimentations du signal dans l'art vidéo. Elle n'étend toutefois pas son analyse à l'art par ordinateur émergent<sup>21</sup>. Nous évoquerons d'autres références au cours de cette recherche, et notamment l'exposition *See This Sound*<sup>22</sup>, qui s'est déroulée en 2009 au Lentos Art Museum à Linz. Elle est un parfait exemple du chemin que nous souhaitons emprunter ici. Un catalogue et deux ouvrages d'« audiovisuologie » proposent des approches multimodales théoriques, esthétiques, archéologiques, scientifiques et perceptives. Son approche singulière illustre une manière transdisciplinaire d'aborder ces formes artistiques. L'un des ouvrages explore différents champs artistiques (le film abstrait, le cinéma étendu, l'art interactif, l'art vidéo, le jeu vidéo, l'art sonore...). Plus récemment, Ana Carvalho et Cornelia Lund – deux théoriciennes de l'image et du son – pointent ce qu'elles nomment un tournant audio-visuel au cours du XX<sup>e</sup> siècle. Dans leur ouvrage *The Audiovisual Breakthrough*<sup>23</sup>, elles identifient cinq grands domaines artistiques historiques et contemporains dans lesquels les artistes ont développé de nouveaux langages fondés sur l'image et le son : la musique visuelle, le cinéma étendu, le *Live Cinema*, le *Vjing*, et la performance audiovisuelle<sup>24</sup>. Pour définir cet ensemble, elles ont interrogé de nombreux artistes pour connaître leurs sentiments d'appartenance à ces mouvances.

18 Terme couramment utilisé au Canada.

19 Exposition *Sons & Lumières. Une histoire du son dans l'art du XX<sup>e</sup> siècle*, Paris, Centre Pompidou, du 22 septembre 2004 au 3 janvier 2005. Commissaires : Sophie Duplaix et Marcella Lista.

20 Des instruments visuels similaires à un piano, mais muets et projetant de la lumière.

21 La relation entre les musiques électroniques et l'art est abordée dans une autre exposition du Centre Pompidou, *Sonic Process*, en 2002, mais avec peu de place donnée à l'image (Exposition *Sonic Process. Une nouvelle géographie des sons*, Paris, Centre Pompidou, du 16 octobre 2002 au 6 janvier 2003. Commissaire : Christine Van Assche).

22 Exposition *See This Sound*, Lentos Art Museum, Linz, du 28 août 2009 au 10 janvier 2010. Commissaire : Cosima Rainer.

23 Ana Carvalho et Cornelia Lund, *The Audiovisual Breakthrough*, Berlin, autoédition, 2015.

24 Nous reviendrons sur ces différentes mouvances au cours de notre recherche.

Dans cet même esprit, pour inspirer et nourrir notre recherche, nous avons réuni un corpus d'œuvres qui traverse les champs de ce que l'on nomme, selon les conventions, les nouveaux média, les arts interactifs, les installations multimédia, l'art numérique, l'art génératif, ou bien encore la performance audiovisuelle. Nous nous concentrerons sur des dispositifs artistiques privilégiant des dialogues entre des écrans (incluant la projection) et des systèmes sonores de diffusion. Le champ restant encore très vaste, nous traiterons de peu d'œuvres relevant du spectacle vivant, du jeu vidéo ou de la réalité virtuelle, tout en notant l'intérêt fort que pourrait susciter l'étude des relations audio-visuelles dans ces formes.

## La difficile circonscription du numérique

L'usage du terme « numérique » nécessite beaucoup de prudence. Il renvoie à une technique — celle de l'encodage de contenus grâce au système binaire — mais apposé au terme image, ou son, il produit un déplacement sujet à ambiguïté. En effet, notre perception est par essence analogique ; tout son ou image numérique doit être converti en analogique (ondes pour le son, lumière pour l'image) pour être vu ou entendu, autrement dit pour être perceptible. La critique littéraire N. Katherine Hayles, spécialiste des humanités numériques<sup>25</sup>, ajoute que les humains développent depuis des milliers d'années « des modifications d'ordre biologique et des prothèses technologiques pour contraindre ces processus analogiques à la numérisation, de l'évolution physiologique nécessaire à la production de la parole jusqu'aux machines numériques les plus sophistiquées<sup>26</sup> ». Pour le philosophe du numérique Bruno Bachimont, le numérique est une « pure idéalité qui n'est abordable qu'à travers des implémentations qui le réalisent et des interfaces qui le “phénoménologisent” (rendent perceptible et interprétable)<sup>27</sup> ». En ce sens, un code numérique, sans dispositif ou processus qui en traduit et en matérialise ce qu'il recèle, est en soi imperceptible. Il n'est qu'un état intermédiaire, qui revêt un caractère insaisissable. C'est ce qui fait dire au philosophe et théoricien de la communication Vilém Flusser que les données numériques appartiennent à

---

25 Les humanités numériques sont l'étude des liens entre les sciences humaines et le numérique.

26 N. Katherine Hayles, *Parole, écriture, code*, Dijon, Les presses du réel, 2015, p. 51.

27 Bruno Bachimont, « Formes, concepts, matières : quels place et rôle pour le numérique et la technique », in Vivien Philizot (dir.), *Technique et design graphique. Outils, médias, savoirs*, Paris, B42, 2020, p. 227.

la catégorie plus large des « non-choses<sup>28</sup> ». Des informations de nature inédite, molles (au sens de *software*), impossibles à saisir avec les mains : « Les images électroniques sur l'écran de télévision, les données stockées dans les ordinateurs, les microfilms et les bobines de film, les hologrammes et les programmes [...] »<sup>29</sup>.

La digitalisation, c'est-à-dire selon Hayles « l'acte de rendre quelque chose discret ou discontinu, soit digital plutôt qu'analogique<sup>30</sup> », est donc un processus central dans l'appréhension de la non-chose numérique. Bruno Bachimont situe le code numérique binaire à un niveau d'abstraction tel qu'il se situe dans un rapport d'indépendance, voire d'indifférence, vis-à-vis du support physique de stockage susceptible de l'accueillir<sup>31</sup>.

C'est ainsi que la discrétisation et le binaire peuvent être pensés, selon Frédérique Vargoz et Emmanuel Guez, comme fournisseurs d'une « clé universelle qui permet de traduire n'importe quel type de données, images, sons, textes, sous la forme d'un code qui peut ensuite être manipulé<sup>32</sup> ». Bachimont décrit comment la manipulation du code permet jusqu'à une commutation des perceptions. Il devient possible pour les artistes d'exploiter les possibilités « gyrovagues<sup>33</sup> » du numérique, par un processus de remanifestation : « [...] une prestation instrumentale est enregistrée, donc numérisée en binaire, puis est relue non pas comme un fichier audio, mais comme un fichier vidéo, qui donne à voir des images à partir du son. On voit donc ce que l'on entend, puisque c'est le même code binaire qui correspond à ce que l'on entend et à ce qu'on voit<sup>34</sup> ». Sur cet aspect, il faut noter que ces types de conversion, comme nous le verrons, ont existé très tôt, dès lors que les images et les sons ont été fixés.

---

28 « Undinge » en allemand.

29 Vilém Flusser, *Choses et non-choses. Esquisses phénoménologiques*, Paris, Jacqueline Chambon, 1996, p. 99.

30 N. Katherine Hayles, *op. cit.*, p. 50.

31 Bruno Bachimont, *op. cit.*, p. 217.

32 Frédérique Vargoz et Emmanuel Guez, « Une histoire de l'ordinateur du point de vue de la théorie des média » in *Cahiers philosophiques* n° 141, « La révolution informatique », Paris, Canopé, mai 2015, p. 62.

33 La définition originale du terme « gyrovague » est celle d'un moine chrétien qui n'était attaché à aucun monastère.

34 Bruno Bachimont, *op. cit.*, p. 226.

En complément de l'approche audio-visuelle, l'approche archéologique des médias propose une pensée non linéaire de l'histoire des médias. Vargoz et Guez constatent qu'il existe « une continuité entre médias analogiques et numériques, l'ordinateur pouvant être considéré, comme chez Kittler<sup>35</sup>, comme l'aboutissement d'un processus engagé depuis la fin du XIX<sup>e</sup> siècle<sup>36</sup> ». L'ordinateur serait le dernier média composite, qui absorbe les anciens (film, radio, télévision, vidéo, téléphone...). Dans ce processus d'ingestion, nous notons que les médias sonores, visuels ou audio-visuels fusionnent dans une écologie numérique — nos smartphones en sont l'exemple type — favorisant, a priori, l'émergence de formes audio-visuelles hybrides. Plutôt qu'une absorption ou une sédimentation, Pierre-Damien Huygues propose quant à lui l'idée de substitution, tout en précisant qu'une nouvelle technique n'est pas systématiquement un gain : « une génération de poussées techniques met de côté, pousse, remplace la précédente. Ce remplacement ne doit pas nécessairement être considéré comme une amélioration. Les techniques ne progressent pas, elles se substituent les unes aux autres<sup>37</sup> ».

La question de l'image numérique — celle du sonore peut se poser dans les mêmes termes — ne doit pas toutefois être ramenée, selon le philosophe Jean Cristofol, qu'à une question technique, mais doit mener à l'analyse, dans la société, de « ses modes de production et de circulation, formes d'appropriation et de transformation<sup>38</sup> ». Les pratiques numériques, pour lui, sont « plurielles, diffuses, transversales, [...] inventent leurs dispositifs en même temps qu'elles inventent des formes, de sorte que les formes sont en même temps des dispositifs<sup>39</sup> ».

En analysant les manières dont se manifeste le numérique dans les œuvres de notre corpus, nous verrons si celui-ci fait émerger de nouvelles relations audio-visuelles.

---

35 Friedrich Kittler était l'un des théoriciens à l'origine de l'archéologie des médias en Allemagne.

36 Frédérique Vargoz et Emmanuel Guez, *op. cit.*, p. 59.

37 Entretien avec Pierre-Damien Huygues, « La technique est-elle sédimentaire ? », in Vivien Philizot, *Technique et design graphique, op. cit.* p. 232.

38 Jean Cristofol, « Le flux et l'image », in Caroline Renard (dir.), *Images numériques ?*, Aix-en-Provence, Presses universitaires de Provence, 2014, p. 264.

39 *Ibid.*

## Depuis le film sonore. Trois sillons d'exploration

Si notre recherche est contemporaine, les liens sont forts entre les développements du pré-cinéma<sup>40</sup>, puis du cinéma qui mettent en jeu une dialectique avec la musique, le bruitage et les dialogues. Déjà en 1920 les formes du cinéma sont multiples : du cinéma d'avant-garde au cinéma d'animation expérimental. À la fin des années 1930, l'émergence du film sonore — dans son acception large d'image animée sonore — constitue le nœud fondamental qui va guider le choix de nos axes de réflexion. À partir de ce moment en effet, le son et l'image peuvent être fixés sur un même support — la pellicule — puis, après montage être diffusés au public, en synchronisation, sans intervention musicale en direct dans la salle.

Nous avons choisi, à partir de cette période-clé (et en élargissant, au fil du XX<sup>e</sup> siècle de l'installation à la performance audiovisuelle), de creuser trois sillons, trois axes signifiants qui constitueront respectivement trois grandes parties : la musique visuelle programmée ; la manipulation numérique des images et des sons saisis du réel ; les manifestations contemporaines du visage et de la voix<sup>41</sup>. Une quatrième et dernière partie documentera notre travail de recherche-crédation qui s'est développé au fil des analyses issues des trois parties.

Le premier axe de travail est celui de ce que l'on a pu appeler la musique visuelle (*Visual Music*), c'est-à-dire un cinéma graphique d'abstraction (et donc non narratif), en conversation avec la musique et le son. Il prend principalement sa source au début des années 1920, quand sont diffusés les trois premiers films abstraits de l'histoire du cinéma : le développement temporel des images y est fortement marqué par le modèle musical. La synthèse optique du son va ensuite permettre le développement d'œuvres totalement novatrices, les premiers films des frères Whitney en étant la quintessence. Les premières images par ordinateur, par analogie, sont aussi des images graphiques abstraites, agencements de formes primitives calculées et générées par l'ordinateur ; d'abord imprimées sur *plotter* ou par impression sur film 16 mm, elles s'inscrivent dans cette tradition, que l'on peut aussi relier à l'art optique. Dans les années 1960, les la-

---

40 Le précinéma (ou pré-cinéma) désigne l'ensemble des procédés, de la lanterne magique aux jouets optiques, qui annoncent les techniques du cinématographe.

41 Il est entendu ces histoires se croisent mais le domaine étant vaste, cette approche directionnelle permet de poser les problématiques essentielles de la question du numérique.

boratoires Bell sont l'un des pôles majeurs d'éclosion de la musique et de l'art par ordinateur. Nous verrons comment s'y sont croisées différentes formes visuelles et sonores par ordinateur, à travers notamment les films de l'artiste Lillian Schwartz. Avec le développement des interfaces graphiques et des langages de programmation, la création numérique se développe et des artistes développent leurs propres outils de création audio-visuelle. Citons, pour le début des années 1990, le film *Moondrum* de John Whitney, créé avec son logiciel de composition audiovisuelle Whitney-Reed RDTD, la suite interactive *Audiovisual Environment Suite (AVES)* de Golan Levin, dans les années 2000, ou bien les œuvres contemporaines audiographiques de Ryoji Ikeda. Parfois tellement facilitée par les outils numériques que vidée de sa substance, la création abstraite retrouve du sens à travers l'utilisation des données. Le collectif Semiconductor utilise l'enregistrement du son d'un sol et des données scientifiques comme matériau pour une installation audiovisuelle. Peter Beyls récupère en temps réel des données éoliennes dans une installation audiographique à mi-chemin entre l'abstraction et la visualisation. Nous analyserons comment ces formes se développent avec le numérique pour devenir des audio-visualisations de nos données.

Le deuxième sillon est celui de la saisie du réel. Cette partie se propose d'étudier ce que devient la saisie des images et des sons du réel, à travers les processus de capture et manipulation transformées par le numérique. Comprendre comment nos images et nos sons sont devenus le cœur de bases de données. D'analyser, également, comment la facilitation de la production d'images et de sons entraîne une prolifération telle que les processus d'automatisation, d'analyse, de recommandation ou de traitement, se substituent à notre discernement. L'être humain n'a parfois plus de prise sur certains processus, qui s'exercent directement entre les machines. D'où un questionnement sur ce que celles-ci peuvent voir et entendre. Chercher à savoir comment se développent la surveillance et les détections automatisées. Enfin, adopter les angles de la matérialité et de l'impact écologique pour explorer la face cachée des infrastructures du numérique. Nous évoquerons des installations audiovisuelles interactives pionnières telles que la *Time Machine* de Piotr Kowalski ou *The Waves* de Thierry Kuntzel, qui matérialisent deux types importants de manipulation de l'axe-temps. L'esthétique des bases de données (dont Lev Manovich est un explorateur) et la question du réemploi

seront traitées à partir des formes audio-visuelles de *Live Cinema* de SLIDERS\_Lab et celle, directement ancrée dans les usages sociaux numériques, de Christian Marclay, qui utilise et mixe des vidéos provenant de l'application Snapchat dans sa série d'installations *Sound Stories*. Les média situés seront illustrés par différentes œuvres dont les *Field\_works* de Masaki Fujihata, tandis qu'avec une installation de Moritz Fehr, *Colosseum*, qui documente l'extraction de métaux destinés à la fabrication d'appareils technologiques, nous aborderons la question de l'écologie et des média.

Enfin, nous chercherons à suivre ce qui était en peinture la figure du portrait, mais qui, avec l'avènement du cinéma parlant, est devenu une notion plus large, depuis le portrait parlant interactif jusqu'aux choralités d'agents conversationnels. Du *Portrait n° 1* de Luc Courchesne — exercice pionnier de vidéo interactive — jusqu'aux montages spatialisés de *Vlogs* de Natalie Bookchin, en passant par des agents artificiels personifiés, nous interrogerons les modalités visuelles et sonores contemporaines du visage et de la voix.

Au début de chaque partie, nous opérerons en premier lieu une archéologie de ce qui constitue selon nous le champ des *arts audiovisuels*, à savoir l'ensemble des formes plastiques qui mêlent l'image en mouvement et le son depuis la naissance du cinéma et de l'animation, en passant par l'art vidéo et le cinéma expérimental, jusqu'aux nouveaux médias. Nous examinerons, au fil de ces différents points de rencontre, comment les natures respectives du son et de l'image ont pu muter au fil des évolutions technologiques, à la fois par ruptures, mais aussi dans une certaine continuité. Nous analyserons ensuite comment ces formes se sont propagées dans l'art, en empruntant les voies de l'interactivité et de la générativité, du temps réel, du traitement des données ou de la connectivité en réseau. Les rôles joués dans l'œuvre d'art par le système, le dispositif, le programme ou l'interface seront au cœur de notre réflexion. Nous nous intéresserons aussi à la perception.

## **Recherche-crédation : des études *audiographiques***

Sur la base des analyses établies dans les trois parties, nous avons mené un travail de recherche-crédation autour de la modélisation et du développement d'un langage-interface destiné à explorer les basiques de ce que pourrait être un art *audiographique*

numérique. En effet, si les possibilités technologiques numériques de manipulation de l'image et du son paraissent à présent infinies, il n'existe pas réellement, à notre sens, de solfège image-son. Un solfège qui, au-delà de celui du cinéma, poserait des bases de relations entre les arts visuels et les arts sonores.

Sur la base de la première partie dédiée à la musique visuelle programmée, nous souhaitons donc créer des séries d'études audio-visuelles expérimentales (sur le modèle du Bauhaus par exemple), dans lesquelles s'exercent différentes combinaisons d'image et de son.

Nous avons en effet constaté que même si des croisements avaient eu lieu dans les années 1960 entre la musique par ordinateur (apparue un peu moins d'une dizaine d'années avant l'image animée par ordinateur) et l'image animée par ordinateur, ce n'est que dans les années 1990 qu'ont émergé en masse des œuvres numériques dans lesquelles l'image graphique synthétique et le son de synthèse peuvent être générés ou manipulés avec la même typologie d'outils numériques. L'infinité des possibilités actuelles des outils ont, nous le pensons, eu pour effet de ne pas permettre une exploration poussée des croisements élémentaires indispensables à une pensée équilibrant l'audio et le graphique. En résumé, nous pensons donc que les bases de la synthèse sonore et ceux de la synthèse graphique sont rarement pensés ensemble, dès l'origine de l'œuvre, aussi nous avons conceptualisé un logiciel permettant d'explorer ces relations audio-graphiques élémentaires. Il s'agit, par exemple, de mettre en scène un point animé et un son, et d'explorer différentes possibilités d'interactions entre les deux composantes. Nous avons pu présenter nos travaux à différentes occasions, et nous avons pu obtenir une bourse destinée à financer le développement d'une application web, *audiographic lab*<sup>42</sup>. Celle-ci permet de manipuler en ligne, en noir et blanc, des formes primaires graphiques et sonores.

Il ne s'agit pas de chercher à élaborer une grammaire universelle, mais de contribuer à faire en sorte que se développe une pensée des relations entre l'image et le son dans le domaine artistique.

---

42 Accessible à l'adresse : <http://www.audiographiclab.com>, consulté le 30/01/23.

## Première partie.

### **Formes de la musique visuelle programmée : de la composition d'abstractions visuelles et sonores à l'info-esthétique**

Dans le champ des formes audio-visuelles contemporaines s'appuyant sur le numérique peuvent être identifiées une large majorité de pièces dont la composante visuelle est abstraite. L'image animée abstraite émerge dans les années 1920, en marge des formes de cinéma qui tirent leur image de la captation du réel. Elle se développe au sein d'un cinéma d'animation plastique qui, dans la lignée de la peinture abstraite, cherche à s'abstraire du réel pour faire naître de nouveaux langages. Ces premiers films abstraits pionniers, d'abord muets — mais leurs structures s'inspirent du modèle musical —, puis musicaux, questionnent à leur manière, comme le cinéma à l'arrivée du sonore, la relation entre l'image animée et les formes musicales.

Nous redonnerons synthétiquement, dans un premier chapitre, les moments-clés de l'histoire de cette musique visuelle, mais nous nous focaliserons surtout sur une série d'artistes qui ont cherché à développer un véritable langage entre le graphique et le musical. C'est notamment le cas des frères Whitney, dans les années 1940. En se positionnant comme auteurs d'image, mais aussi de son, ils inventent une machine d'animation dédiée à servir leur projet conceptuel : faire émerger une forme de musique audio-visuelle inédite.

Après-guerre, parallèlement à l'évolution de l'animation expérimentale et de l'art vidéo, les relations entre image graphique et son vont se développer au sein d'une branche émergente : l'art par ordinateur. Le chercheur Bernard Caillaud divise l'histoire de cet art par ordinateur en trois périodes<sup>43</sup>. La première, de 1960 à 1965, est celle de l'art « construit », (en référence au constructivisme), utilisant des formes primaires, l'aléatoire, et presque uniquement le noir et blanc. La deuxième période (1976-1985),

---

<sup>43</sup> Bernard Caillaud, *La création numérique visuelle, Aspects du Computer Art depuis ses origines*, Paris, Euro-pia, 2001, p. 5.

âge des micro-ordinateurs et des terminaux graphiques, voit renaître dans l'industrie les grandes ambitions de simulation du réel, occultant alors les pionniers du « computer art ». Enfin, la troisième et dernière période (1986-2000), est celle de l'explosion du numérique et de la montée en puissance des matériels. Si cette structuration illustre bien une évolution de l'art visuel par ordinateur, elle ne traite pas, comme beaucoup d'autres histoires de l'art par ordinateur, du sonore et de la musique informatique<sup>44</sup>. Pour cela il est intéressant de relire, de manière croisée, les histoires dans lesquelles les sons et les images sont devenus des entités programmées et programmables et de chercher les moments où des artistes ont pu expérimenter avec les deux matériaux. Dans le deuxième chapitre, c'est donc vers 1957<sup>45</sup>, avec l'émergence de la musique informatique aux laboratoires Bell, — c'est-à-dire une dizaine d'années précédant les premières expérimentations visuelles programmées — que nous débuterons une analyse visant à examiner les conditions de création des premiers films programmés et leur relation à la musique ou au son. Nous montrerons que même si la musique informatique et l'art visuel par ordinateur ont cohabité, seul un petit nombre d'artistes, de John Stehura à Lillian Schwartz, ont réellement tenté d'établir des ponts entre ces deux champs.

Dans un troisième chapitre, nous évoquerons le retour de l'ordinateur à l'ère de l'installation du numérique dans notre société, à travers les notions de programme, d'interface et de temps réel. Comme le précise Darko Fritz, spécialiste de cette période pionnière, le développement combiné au milieu des années 1970 de l'infographie, de la recherche et des effets spéciaux réalistes, amorce leur entrée dans les mondes de l'industrie et du domaine militaire. Cela a conduit « à un rejet de l'art par ordinateur et à son exclusion presque totale de la scène de l'art contemporain<sup>46</sup> ». Les artistes se dirigent alors plutôt vers « l'art vidéo, l'art conceptuel et l'art non objet [...] »<sup>47</sup>. L'avènement du micro-ordinateur individuel couplé au développement des interfaces gra-

---

44 Parmi les exceptions, notons toutefois qu'Abraham Moles, dans son ouvrage *Art et ordinateur* de 1971, consacre déjà tout un chapitre à la musique (Abraham A. Moles, *Art et ordinateur*, Paris, Blusson, 1990).

45 Quelques expérimentations pré-existent, mais l'émergence de l'informatique musicale aux laboratoires Bell reste un moment clé.

46 Darko Fritz, « La notion de "programme" dans l'art des années 1960 — art concret, art par ordinateur et art conceptuel », in David-Olivier Lartigaud (dir.), *Art++*, Orléans, HXX, coll « Script », 2011, p. 38.

47 *Ibid.*, p. 38.

phiques et des logiciels offre de nouvelles possibilités d'interaction entre image et son. Nous chercherons à montrer que ces nouvelles potentialités ont submergé la nécessité de penser plus en profondeur les relations entre image graphique et son, et mis de côté la potentielle mise en place de solfèges audio-graphiques, au profit d'effets visuels et sonores trop systématiques.

Dans le quatrième et dernier chapitre, nous examinerons les mutations entre arts et sciences à travers l'utilisation des données scientifiques, qui nous révèlent des parts cachées de nos environnements. La production intensive de données a donné lieu à des collaborations entre artistes et scientifiques, les premiers cherchant à matérialiser de manière singulière le matériau des seconds. D'autres artistes utilisent des données physiques provenant de capteurs, ou bien produisent leurs propres données selon des processus algorithmiques génératifs. Celles-ci sont ensuite sonifiées et/ou visualisées.

## 1.1 *Abstractions graphiques et musicales du début du XX<sup>e</sup> siècle*

### 1.1.1 De la musique visuelle



Fig. 1 : Paul Klee, *Fuge in Rot* (1921)  
© Media Art Net

On attribue communément la paternité du terme « musique visuelle » au critique Roger Fry. En 1912, celui-ci emploie l'expression « Visual Music » afin de caractériser

un ensemble, présenté dans le cadre d'une exposition, de travaux picturaux s'éloignant de la figuration<sup>48</sup>. Ce terme, peu usité pendant un grand laps de temps, n'a été repris que dans les années 1980 afin de décrire certains films abstraits des années 1930<sup>49</sup>. Pourtant au départ, il n'évoque que la musicalité d'une composition visuelle — fixe qui plus est —, et sous-tend en conséquence une absence de musique. Son acception s'est progressivement étendue à des formes abstraites, animées et sonores. Le terme est ainsi resté et englobe aujourd'hui différentes formes audiovisuelles d'art, numériques ou non, s'inscrivant dans l'histoire des films abstraits musicaux.

Avant même ce concept de musique visuelle, des correspondances entre musique et visuel ont été imaginées. Dans *Musique et arts plastiques*<sup>50</sup>, Jean-Yves Bosseur raconte que déjà au XVI<sup>e</sup> siècle, le peintre Giuseppe Arcimboldo avait travaillé sur un système d'équivalences entre hauteurs de son et graduation de gris, ou bien qu'en 1740 l'Abbé Louis-Bertrand Castel avait mis en place un tableau de correspondances entre couleurs et notes, dans le but de développer un « clavecin pour les yeux » (appelé aussi clavecin oculaire). Cette création n'a jamais été finalisée, mais elle avait pour ambition de proposer des projections spectrales grâce à un clavier à touches qui aurait activé des lammelles de tissus teintés passant devant une flamme<sup>51</sup>. Le compositeur Scriabine, qui avait consulté les traités de Castel, était lui aussi à la recherche d'un art total convoquant tous les sens, et avait établi dans sa partition de *Prométhée* (1915) des correspondances entre notes et couleurs. Tout un courant est issu de ces mises en relation entre couleur et musique, il se nomme « Color Music » (Musique des couleurs). Il va principalement concerner des performances liées à la projection de lumière<sup>52</sup>.

48 Il est souvent écrit que Fry ciblait le travail de Kandinsky, mais l'artiste Joost Rekveld, à la lumière d'une nouvelle exploration de ces documents, montre que la toute première utilisation du terme par Fry en 1912 ciblait les œuvres d'une exposition dans laquelle Kandinsky ne figurait pas, mais on pouvait y voir des tableaux post-impressionnistes, de Picasso, Cézanne, Matisse, ou bien encore de Duncan Grant (Joost Rekveld, « The origin of the term Visual Music », 2013, <http://www.joostrekveld.net/?p=1105>, consulté le 30/01/2023).

49 Notamment ceux d'Oskar Fischinger, analysés par William Moritz. Nous en parlerons dans la section suivante.

50 Jean-Yves Bosseur, *Musique et Arts Plastiques, Interactions au XX<sup>e</sup> siècle*, Paris, Minerve, coll. « Musique Ouverte », 2006, p. 9.

51 Sophie Dupleix, Marcella Lista (dir.), *Sons et lumières. Une histoire du son dans l'art du XX<sup>e</sup> siècle*, Paris, Centre Pompidou, 2004, p. 31.

52 Un certain nombre d'orgues à lumière vont donc être inventés. Citons l'orgue à couleurs d'Alexander Wallace Rimington et celui de Brainbridge Bishop en 1893, ou bien le piano optophonique de de Baranoff Rossiné en 1915. Souvent muets (même si celui de Rimington, par exemple, produisait à la fois son et lumière), ils servaient à accompagner la musique. L'un des plus intéressants à ce titre est le Clavilux de Thomas Wilfred, à partir de 1922. Cet instrument comporte une console qui déclenche et contrôle les rotations par cran (une centaine) de disques peints au sein d'un dispositif composé de

Dans le domaine de la peinture, plusieurs mouvements (comme le synchronisme et le mouvement musicaliste) vont aussi mettre en œuvre des principes de correspondance. L'artiste le plus connu à avoir travaillé en ce sens est probablement Vassily Kandinsky. Il tâche d'établir des associations entre couleur, timbre et émotion (résonance intérieure). Mais le développement de ce type d'associations va vite s'épuiser. D'autres peintres vont plutôt se diriger vers des emprunts structurels à la musique, qui les mèneront vers un registre formes picturales inédites. La musique de Bach, notamment, est réexaminée à la lumière de ces nouvelles orientations dans la peinture et retrouve une modernité, d'abord de manière superficielle par la simple résurgence d'emprunts terminologiques pour les titres de peintures (une utilisation métaphorique en somme), puis, peu à peu, par l'emploi de techniques ou de théories de composition musicale. Par exemple, la forme musicale de la fugue<sup>53</sup>, habile usage du contrepoint et de la polyphonie, peut par analogie s'appliquer au processus de peinture. Dans un article sur Paul Klee consacré à l'utilisation de la fugue chez les peintres d'avant-garde, Peter Vergo rappelle que celle-ci « offre des possibilités presque illimitées de déployer des processus élaborés de transformation des thèmes ou de motifs. Une fois exposé, le sujet de la fugue [...] peut être réduit ou étendu, c'est-à-dire allongé ou raccourci simplement en utilisant des valeurs de notes plus longues ou brèves. On peut y apporter des variations rythmiques, le renverser et le jouer à l'envers ou en sens contraire<sup>54</sup> ». Klee développera, au sein du Bauhaus, un langage formel moderne reposant sur des éléments fondamentaux, dont une grande partie est issue de réflexions sur les relations structurelles entre musique et peinture.

---

prismes mobiles et de lentilles. Wilfred proposait, grâce à cet instrument, des performances au cours desquelles il pouvait en direct manipuler la lumière. Il nommait « Lumia » cette forme d'art, et, contrairement à d'autres inventeurs d'orgue à couleur, il ne faisait pas d'analogies avec la musique ; il s'agissait bel et bien d'un art, exclusif, de la lumière. Il semblait en revanche avoir deux méthodes pour combiner ses lumières avec de la musique : soit il commandait un accompagnement musical pensé pour une œuvre de lumière existante. Soit l'œuvre de lumière était composée comme un accompagnement visuel pour une œuvre musicale donnée (Sophie Duplex, Marcella Lista [dir.], *op. cit.*, p. 192).

53 Voici la définition qu'en donne André Hodeir : « La fugue (fuga : fuite) est une composition de style contrapunctique, qui se fonde sur l'usage de l'imitation et la prépondérance d'un thème générateur, court, mais caractéristique, nommé sujet. Apparue vers la fin du XVII<sup>e</sup> siècle, elle se rattache à la tradition polyphonique des XV<sup>e</sup> et XVI<sup>e</sup>, dont elle est en quelque sorte le couronnement : dans nulle autre forme l'équivalence des différentes voix n'est aussi évidente ; nulle autre forme n'atteint à une unité aussi parfaite » (André Hodeir, *Les formes de la musique*, Paris, PUF, 2003, p. 52).

54 Peter Vergo, « Bach et la fugue comme modèle pour les peintres d'avant-garde », in Marcella Lista (dir.), *Paul Klee, Polyphonies*, catalogue d'exposition, Paris, Actes Sud/Cité de la Musique, 2011, p. 55.

### 1.1.2 Naissance d'un cinéma graphique musical

La période des années 1920 est extrêmement riche pour l'image animée. Le cinéma s'y développe et trouve par le montage de nouveaux langages qui donnent naissance à des formes cinématographiques d'avant-garde. Dans cette effervescence des peintres, s'inspirant du modèle musical, vont explorer la forme animée.

#### Des formes animées à la musique. Le premier film abstrait musical



Fig. 2 : Walter Ruttmann, *Opus 1* (1921)  
© See This Sound

En 1919, Walter Ruttmann imagine une nouvelle forme d'art, qui se situe « aux franges de la peinture et de la musique<sup>55</sup> » ; le développement temporel en constitue l'essence formelle. Il met en pratique ses théories deux ans plus tard et devient l'un des pionniers de ce courant émergent avec son film *Lichtspiel Op.1* (1921)<sup>56</sup>. Il est l'un des trois premiers films abstraits de l'histoire de l'art. Les deux autres, sortis la même année, sont *Symphonie Diagonale* (1921)<sup>57</sup> de Viking Eggeling et *Rythme 21* (1921)<sup>58</sup> de Hans Richter.

55 Walter Ruttmann, « Peindre avec le temps », trad. A. Pernet, in Yann Beauvais (dir.), *Musique Film*, Paris, Scratch/Cinémathèque française, 1986, p. 18.

56 Une version est disponible en ligne sur Media Art Net, <http://www.medienkunstnetz.de/works/opus1/video/1/>, consulté le 30/01/23.

57 Une version est disponible en ligne sur le site du festival Punto y Raya, <https://vimeo.com/576389010>, consulté le 30/01/23.

58 Une version est disponible en ligne sur le site du festival Punto y Raya, <https://vimeo.com/262035970>, consulté le 30/01/23.



Fig. 3 : Viking Eggeling, *Symphonie diagonale* (1921-1924)  
© RE:VOIR Vidéo

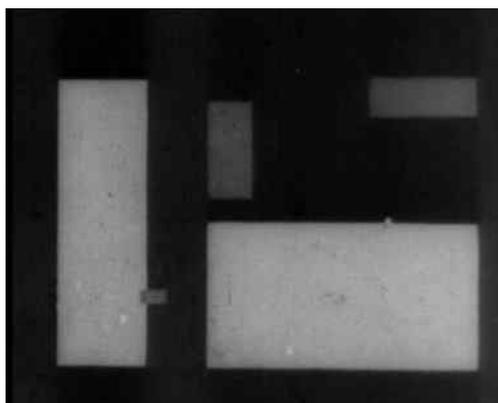


Fig. 4 : Hans Richter, *Rythme 21* (1921-1924)  
© RE:VOIR Vidéo

L'histoire de ces deux films, muets et noir et blanc, est désormais très étudiée<sup>59</sup>. Rappelons surtout que les deux artistes ont d'abord travaillé ensemble sur de longs rouleaux. Ils dessinaient sur l'horizontalité des formes, chacune représentant un état de ce qu'ils imaginaient devenir une séquence animée. Leur ambition était alors de créer un langage universel<sup>60</sup>. Eggeling, qui était proche du mouvement De Stijl<sup>61</sup>, a poursuivi en ce sens en faisant évoluer des compositions abstraites très graphiques dans son film *Symphonie Diagonale* (1921-1924). Chaque motif constitutif s'apparente, chez Eggeling, à un instrument. Richter, en rupture avec la conception d'Eggeling, s'attache à la question du rythme et travaille seul sur son premier film *Rythme 21* (1921), entièrement constitué d'apparitions et de disparitions de rectangles blancs ou gris sur fond noir. Il crée des jeux dans le rapport entre figure et fond. C'est, avant l'heure, un film structurel qui questionne des fondamentaux du cinéma. Celui de Viking Eggeling, annonçant le motion design, reste aussi encore d'actualité.

59 Citons par exemple le catalogue d'une exposition monographique sur Hans Richter en 2013 au Centre Pompidou Metz, et dont le commissaire était Philippe-Alain Michaud (Philippe-Alain Michaud, Timothy O. Benson [dir.], *Hans Richter, la traversée du siècle*, Centre Pompidou-Metz, 2013).

60 Un langage d'abstraction visuelle, qu'Eggeling explicite par des concepts tels que la « Basse générale de la peinture ». Hans Richter précise que tout comme Eggeling, « [...] il était arrivé à sa théorie par le chemin de la musique, et expliquait sa théorie toujours en termes musicaux » (Hans Richter, *Dada, art et anti-art*, Bruxelles, La connaissance, 1965, p. 60).

61 Mouvement fondé par Theo van Doesburg en 1917. Il prônait l'idée d'un art concret, totalement détaché de la représentation figurative : « Le rythme accéléré de la vie contemporaine ayant supprimé peu à peu tous les intervalles, les formes élémentaires furent obligées d'apparaître. La ligne droite répond à la vitesse de la circulation moderne, les plans horizontaux et verticaux aux manipulations minimales, aux plus simples des fonctions de la vie et de la technique industrielle. » (Theo Van Doesburg, « L'élémentarisme et son origine », publié dans la revue *De Stijl* en 1928, in Frédéric Migayrou [dir.], *De Stijl. 1917-1931*, Paris, Centre Pompidou, 2010, p. 263.).

Si les deux réalisateurs s'inspirent de la musique (à travers la notion d'instrument chez Eggeling, par le rythme chez Richter), ces deux films sont pourtant silencieux. Le film abstrait de Ruttmann a en revanche été pensé pour être mis en musique<sup>62</sup>. À cette époque, même pour les films muets classiques, les mises en musique jouées par un orchestre en direct et synchronisées précisément, étaient très rares. Pourtant dans cette œuvre, il existe une partition classique sur laquelle Ruttmann a dessiné des images du film, aux moments-clé correspondants. La musique a ainsi été conçue à partir de la typologie et du mouvement des formes. Dans sa thèse, Carlos Lopez Charles a étudié en détail la composition audio-visuelle<sup>63</sup> et montre bien les rapports serrés entre l'évolution des formes graphiques et la structure musicale qui s'appuie sur celles-ci.

Les trois films ont été diffusés au début des années 1920 dans différentes salles, ce qui leur a permis d'acquérir une certaine notoriété. S'ils sont fondés sur des techniques différentes (de la peinture à la manipulation de rectangles de carton), ils s'inscrivent tout de même dans la veine commune d'un cinéma graphique, défini ainsi par Dominique Willoughby :

Le cinéma graphique se définit par un phénomène spécifique, fondé sur une méthode ou un principe technique commun, à savoir la synthèse du mouvement image par image, chacune de ces images étant produite, inscrite, par quelque méthode graphique, qu'elle soit manuelle ou issue d'un dispositif de traçage ou d'impression. Images et mouvements sont donc construits graphiquement, contrairement aux méthodes d'enregistrement sur le vif, les prises de vues au moyen de caméras chronophotographiques, qui apparaissent un demi-siècle plus tard<sup>64</sup>.

En effet, Willoughby complète sa réflexion en ajoutant qu'en se croisant avec les techniques photographiques, le cinéma graphique étend son acception. Il devient le cinéma des images animées « tant que les mouvements sont perçus comme artéfact<sup>65</sup> ». Se-

62 Une musique a été commandée par Ruttmann au compositeur Max Butting.

63 Carlos Lopez Charles, *La convergence musicale entre les médias visuels et musicaux dans la création de la musique visuelle*, sous la direction de Anne Sédès, École doctorale Esthétique, sciences et technologie des arts, Université de Paris 8, Saint-Denis, 2015.

64 Dominique Willoughby, *Le cinéma graphique. Une histoire des dessins animés : des jouets d'optique au cinéma numérique*, Paris, Textuel, 2009, p. 18.

65 *Ibid.*, p. 27.

lon lui, si l'image abstraite est photographique, et que le mouvement résultant du passage des images est fabriqué de toute pièce (comme dans le cas du *Stop Motion*), on peut alors parler de « cinéma graphique ». Willoughby relève aussi l'utilisation de deux autres termes : celui d'« animation artificielle », par Auguste Pawlowski, et celui de « film synthétique<sup>66</sup> », par Alexandre Alexïeff. La musique visuelle, suivant ces principes, serait donc un cinéma graphique non représentatif, c'est-à-dire un cinéma graphique abstrait.

### De la musique aux formes animées

L'un des spectateurs de ces films fut le peintre Oskar Fischinger. Celui-ci, au début des années 1930, explore ces liens entre musique et mouvement pictural, qui constitueront le point central de son travail. Après de nombreuses expérimentations (avec de la cire par exemple), il va s'appuyer sur des extraits de musique classique pour faire vivre des formes abstraites. Il ne s'agit pas pour lui d'une simple illustration graphique de la musique, mais de développer une expression singulière qui opère avec la musique, qui agit comme un guide.

L'utilisation de la couleur caractérise ses peintures et ses premiers films, mais il a également réalisé dans les années 1930 une série d'études en noir et blanc<sup>67</sup>, dans un mode mettant en rapport figure (des formes graphiques animées blanches) et fond (un simple fond noir). Dans celles-ci on comprend très bien comment les formes se groupent, se séparent, disparaissent, reviennent en fonction des différents mouvements musicaux<sup>68</sup>.

---

66 *Ibid.*, p. 29.

67 Un extrait de *Studie Nr 8* (1931) est disponible sur le site du Center for Visual Music, <https://vimeo.com/35735682>, consulté le 30/01/23.

68 L'une des études (la n<sup>o</sup> 8) utilise le poème symphonique *l'Apprenti sorcier* (1897) de Paul Dukas, utilisé plus tard dans le dessin animé musical *Fantasia* (1940). Fischinger a en effet travaillé un petit temps pour Disney, sans que cela n'aboutisse.

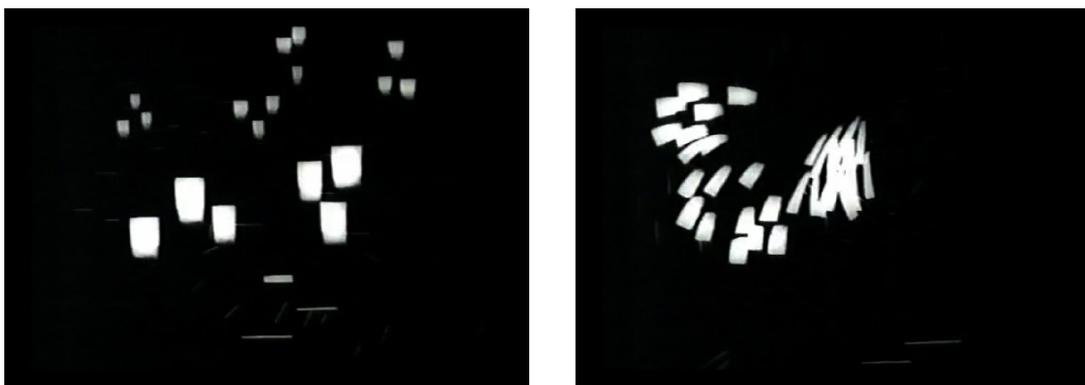


Fig. 5 : Oskar Fischinger, *Studie Nr 8* (1931)  
© Center for Visual Music

Une œuvre postérieure, *Optical Poem* (1938)<sup>69</sup>, dans laquelle se synthétisent ses expérimentations sur une musique de Brahms (la *Danse Hongroise n° 2*) est tout à fait éloquentes ; il y orchestre un jeu équilibré et maîtrisé entre formes graphiques et formes musicales.

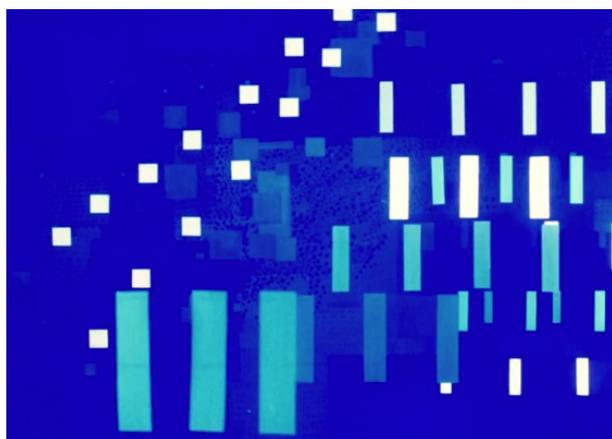


Fig. 6 : Oskar Fischinger, *Optical Poem* (1938)  
© Mubi

Fischinger est probablement l'un des premiers cinéastes d'animation qui utilise la musique comme point de départ pour travailler sur le mouvement. Il invente en quelque sorte la forme du vidéo-clip, dans laquelle le visuel s'appuie sur une musique préexistante. Le théoricien Laurent Jullier<sup>70</sup> rappelle que les critiques ont souvent employé

69 Une bande annonce du film est disponible sur Mubi, <https://mubi.com/fr/films/an-optical-poem/trailer>, consulté le 30/01/2023.

70 Laurent Jullier, Julien Péquignot, « L'effet-clip au cinéma », Kinephanos, 2013, <https://www.kinephanos.ca/2013/effet-clip/>, consulté le 30/01/23.

deux termes descriptifs qui sont restés : celui d'une part de l'effet-clip (ou effet-ballet), situation dans lequel la bande-son dicte l'organisation de l'image — c'est le cas par essence de la forme vidéoclip, dont la composante sonore est un morceau de musique préexistant, ce qui induit d'emblée une relation déséquilibrée, d'asservissement de l'image au son ; d'autre part celui de l'effet-cirque, effet produit par l'orchestre de cirque qui se conforme en permanence à l'action — un saut d'animal ou l'entrée d'un personnage, dont l'un des cas extrêmes est le *Mickeymousing*, c'est-à-dire la sonorisation systématique du moindre mouvement dans les dessins animés.

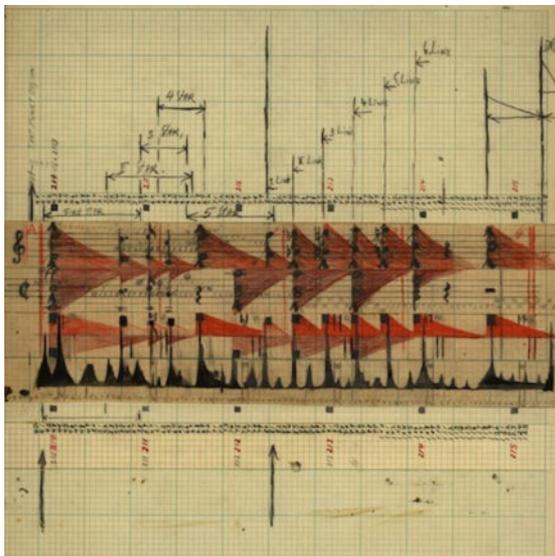


Fig. 7 : Oskar Fischinger, croquis de travail sur une partition  
© Center for Visual Music

Le travail de Fischinger sur les relations entre musique et formes visuelles est un savant mélange de précision et d'indépendance. Il crée des groupes de formes graphiques (similaires à des essaims) qui semblent se caler sur la musique, tout en préservant des libertés d'agencement dans leurs mouvements individuels.

Les films de Fischinger, notamment ses études en noir et blanc, permettent d'explorer le fonctionnement de notre perception face à l'association de mouvements musicaux et de certains groupes de formes. Nous reparlerons de ces principes perceptifs dans la quatrième partie, consacrée à notre recherche-crédation.

### 1.1.3 Du graphique au sonore. La synthèse sonore optique ou la réversibilité d'un système

Au début des années 1930, le cinéma sonore a pu émerger grâce à la stabilisation d'un système de fixation synchronisée de l'image et du son : la pellicule optique. Pour la première fois le son, comme l'image, était présent directement sur la pellicule, sous la forme d'un signal d'amplitude.

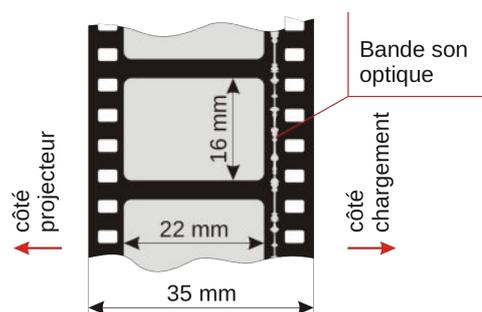


Fig. 8 : Piste optique sur film 35 mm  
Le motif graphique est le signal d'amplitude.  
©Wikipedia

Cette évolution du support a très vite donné lieu aux premières interventions plastiques sur la pellicule, fondatrices du cinéma expérimental<sup>71</sup>, mais a également posé la question du son. Un son étant présent sur la pellicule en tant que représentation graphique, que se passe-t-il si, à l'inverse, on produit une représentation graphique que l'on expose sur la pellicule sur la partie réservée au son ?

À la même période, au début des années 1930, trois pôles d'expérimentation se développent. L'ingénieur Rudolf Pfenninger, le cinéaste Oskar Fischinger et des cinéastes d'animation soviétiques en URSS vont découvrir qu'en dessinant des motifs graphiques réguliers sur la partie de la pellicule vierge dédiée au son, il résultera au passage dans l'appareil de projection (et de diffusion sonore) des sons au timbre étrange, totalement synthétiques. La paternité de cette invention est en général accordée à Rudolf Emil Pfenninger, qui a développé le système le plus stabilisé en premier,

<sup>71</sup> Par exemple le cinéaste d'animation Len Lye dans *Colour Box* (1935). Également le lettriste Isidore Isou avec son *Traité de bave et d'éternité* (1951). L'aspect sonore de ce film est d'ailleurs tout aussi intéressant. Pour qualifier la relation entre l'image et le son dans ce film, Isou utilise le terme « disrépance ». Celui-ci définit une disjonction totale de l'image et du son, la bande-son et l'image étant totalement autonomes.

comme il le montre dans le film documentaire *Tönende Handschrift* [L'Écriture sonore] en 1932. Pfenninger décrivait ces types de son comme « venus de nulle part ».

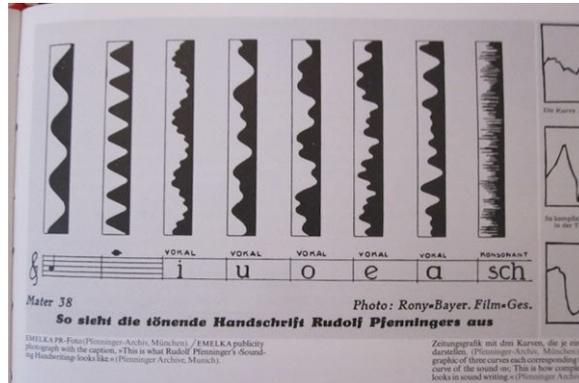


Fig. 9 : Rudolf Pfenninger et ses motifs graphiques  
@ Centre Pompidou

Il s'agit, à l'heure où les seuls instruments singuliers étaient le Theremin et les ondes Martenot, des premiers sons synthétiques, ancêtres des sons de synthétiseur.



Fig. 10 : Oskar Fischinger et ses motifs graphiques  
Il s'agit juste d'une pose à vue publicitaire, pour ne pas dévoiler le principe.  
© Media Art Net

D'autres expérimentations similaires ont aussi lieu en URSS à cette époque<sup>72</sup>, mais citons ici l'exemple d'une œuvre d'Oskar Fischinger, très emblématique de ce procédé,

<sup>72</sup> Voir pour cela l'ouvrage d'Andrei Smirnov, *Sound in Z* (Andrei Smirnov, *Sound in Z: Experiments in Sound and Electronic Music in Early 20th-century Russia*, Köln, Walther König, 2013).

qu'il aurait expérimenté en même temps que Pfenninger. Il s'agit d'une compilation d'expérimentations sur ces formes graphiques destinées à être « sonifiées » qu'il nomme « ornements sonores » : *Tönende Ornamente* [Ornements sonores] (1932). Il les décrit ainsi : « motifs dessinés sur le papier à la plume et à l'encre photographiés directement sur le bord de la pellicule réservé à la bande-son<sup>73</sup> ». Philippe Langlois s'y intéresse particulièrement :

Effectivement, sur le côté gauche de l'image, défile verticalement, de bas en haut, le dessin synchronisé de l'onde sonore qui est entendue. Le long du dessin de ces ondes sonores en mouvement apparaissent des petits rectangles à chaque nouvelle émission sonore, indiquant la hauteur de la note produite dans la notation internationale [...]. *Tönende Ornamente* nous présente ainsi l'étendue spectrale du son synthétique. Au début du film, une gamme descendante qui explore le registre grave, suivie d'une déambulation en octave dans la partie aiguë du spectre sonore. Fischinger présente également les différents timbres et types d'attaque que l'on peut obtenir avec les sons optiques<sup>74</sup>.

Fischinger étant l'un des pionniers du film abstrait, passionné de musique et maître d'une technique de son synthétique, il peut paraître étonnant qu'il n'ait pas réalisé de film abstrait dont la composante sonore eût été entièrement synthétique.

Ce détournement de la bande son optique sera ensuite repris par de nombreux cinéastes expérimentaux<sup>75</sup>, mais va, dès les années 1940, ouvrir la voie à un langage véritablement audio-visuel.

---

73 Thomas Y. Levin, « Des sons venus de nulle part. Rudolf Pfenninger et l'archéologie du son synthétique », in Sophie Duplaix, Marcella Lista (dir.), *op. cit.*, p. 53.

74 Philippe Langlois, « Oskar Fischinger », OLATS, 2001, <http://www.olats.org/pionniers/pp/fischinger/fischinger.php>, consulté le 08/06/2022.

75 Parmi eux le cinéaste d'animation canadien Norman McLaren, qui dessinera lui aussi directement du son sur la pellicule, puis systématisera un système de cartes sur lesquels sont dessinés des signaux graphiques afin de créer ce qu'il nomme du « son animé » de manière extrêmement précise. Le principe du son optique synthétique est au cœur de l'installation de cinéma étendu *Light Music* (1975) de Lis Rhodes, qui montre en projection le signal qui génère le son, créant une mise en abyme.

### 1.1.4 Un rééquilibrage de l'image et du son. La musique audio-visuelle des Frères Whitney

À l'orée des années 1940, plusieurs typologies de films abstraits — dans leur relation à la musique — ont donc vu le jour. Les films muets, *Symphonie diagonale* de Viking Eggeling et *Rythme 21* de Hans Richter, se fondent sur des développements visuels s'inspirant du contrepoint musical. *Opus 1* de Walter Ruttmann, également muet, est toutefois conçu pour être accompagné par une bande musicale originale, classique et fidèle à l'époque, commandée à un compositeur. Ici, c'est l'image, la peinture animée, qui pose de nouvelles bases. Oskar Fischinger, de son côté, a continué de creuser le sillon de l'animation graphique, mais en s'appuyant sur des pièces préexistantes de musique classique.

Au début des années 1940, deux frères, américains, découvrent tous ces films. L'un d'eux, John Whitney, a déjà depuis quelques années des velléités musicales et de création visuelle. Au Pomona College, à Claremont en Californie, il s'imagine déjà compositeur quand, au même moment, la caméra, toute nouvelle à l'époque, suscite sa curiosité. Après deux ans d'étude, il part à Paris pour un an, juste avant la guerre en 1939. Son voisin René Liebowitz, futur chef d'orchestre renommé et jeune étudiant brillant de Schönberg, lui enseigne la composition sérielle tout juste naissante. Celle-ci est alors considérée comme avant-gardiste<sup>76</sup>. En parallèle, comme il possède une caméra 8 mm, Whitney commence à envisager cet appareil comme un outil créatif qui lui permettrait de développer une forme de film « abstrait », dont la structure s'inspirerait de la musique.

En découvrant les films de Fischinger, les deux frères Whitney se posent une question simple, mais fondamentale. Pourquoi ce cinéaste utilise-t-il des musiques préexistantes ? Pourquoi l'artiste ne pourrait-il pas créer à la fois le visuel, et la musique ? La question est d'autant plus pertinente qu'Oskar Fischinger vient de mettre au point une technique de synthèse sonore tout à fait novatrice (nous avons parlé à la section pré-

---

<sup>76</sup> John Whitney, *A personal search : for the complementary of Music and Visual Art*, 1992, DVD Pyramid Media.

cédente). C'est cette technique sonore nouvelle, le son optique synthétique, qui va permettre aux frères Whitney de se départir du domaine musical traditionnel pour affirmer une nouvelle façon d'envisager les relations entre image abstraite et création sonore, en fabriquant synthétiquement à la fois l'image, mais aussi le son.

### ***Five Abstract Film Exercises***

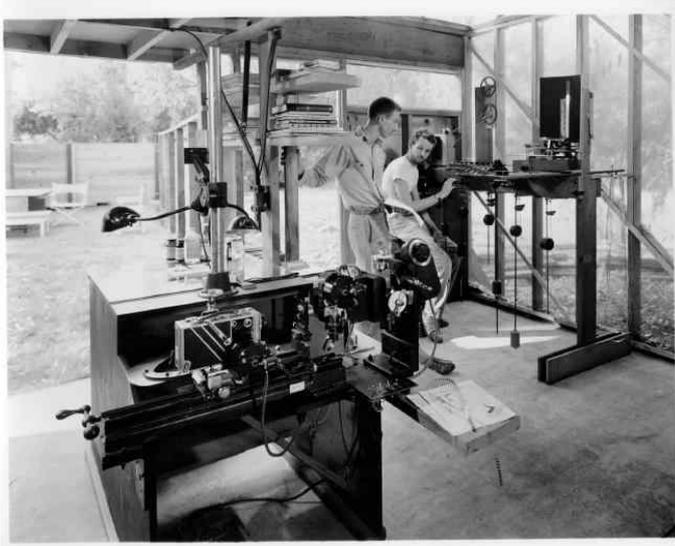


Fig. 11 : Les frères Whitney dans leur atelier  
© Robert Russett, Cecile Starr (dir.), *Experimental Animation. Origins of a New Art*, Cambridge, Da Capo Press, 1976

Leur premier film n'est constitué que d'un simple disque et d'un rectangle. Pour travailler sur l'animation de ces formes, ils réalisent des cartes avec encoches, ainsi qu'une série de petits pochoirs représentant divers états de la relation de superposition entre les deux formes, en jouant sur les contreformes. Les deux frères sont convaincus qu'ils doivent travailler simultanément l'image et le son. John Whitney met alors au point un système constitué de pendules autonomes dont l'oscillation peut être contrôlée très précisément.

Notre instrument à son subsonique consistait en une série de pendules reliés mécaniquement à un coin optique. La fonction de ce coin était la même que celle du modulateur de lumière dans les enregistrements optiques de son que l'on utilise habituellement pour le cinéma. L'instrument ne générait aucun son audible. Au lieu de cela, une bande-son optique de dimension standard était

synthétiquement exposée sur le film ; après traitement, elle pouvait être lue à l'aide d'un projecteur de cinéma ordinaire<sup>77</sup>.

Le résultat tangible de toutes ces expérimentations est un ensemble de cinq films courts, abstraits et sonores, intitulés *Five Abstract Film Exercises*<sup>78</sup> (1944-45). À notre connaissance, ils sont les premiers films véritablement sonores et graphiques, dans lesquels l'image et le son sont générés synthétiquement, et dont l'équilibre entre ces deux composantes est parfait, au sens où l'image ne prédomine pas sur le son et inversement. Il est difficile de dire s'il s'agit de films abstraits sonores, ou de musiques illustrées. John Whitney parlera de « musique audiovisuelle ». L'usage de sons épurés et de formes simples favorise cette sensation, comme le note le critique William Moritz :

À l'incandescence mystérieuse et sensuelle de ces formes au néon répondent des trames sonores représentant l'enfance de l'électronique créées par les frères Whitney directement sur la piste sonore à l'aide d'un étalonnage minutieux. Créés avant le perfectionnement de la bande-sonore, ces sons — exotiques en raison de leur pureté et répartis en glissandos et en échos de pulsations rigoureusement mathématiques — avaient à l'époque quelque chose de révolutionnaire et de choquant<sup>79</sup>.

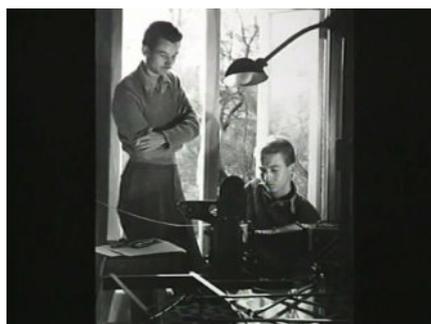
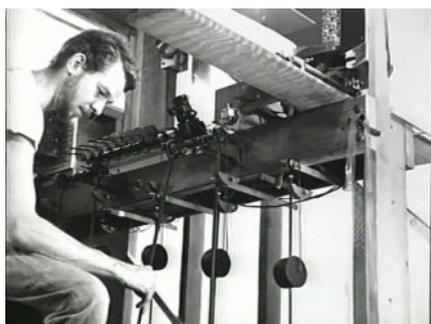


Fig. 12 : À gauche le pendule, à droite les deux frères  
© Robert Russett, Cecile Starr (dir.), *Experimental Animation. Origins of a New Art*, Cambridge, Da Capo Press, 1976

77 John Whitney, « Moving pictures and electronic music » [Images en mouvement et musique électronique], trad. Jean-François Allain, in Sophie Duplaix, Marcella Lista (dir.), *op. cit.*, p. 213.

78 Série aussi appelée *Five Film Exercises*.

79 Texte de William Moritz sur le site de Light Cone, distributeur français des films.  
<https://lightcone.org/en/film-11409-film-exercise-n5>, consulté le 30/01/23.

L'aspect visuel de ces films, basé sur les mouvements de surimpressions de formes géométriques de couleur vive (similaire à des néons) est très futuriste pour l'époque, mais le plus étonnant réside probablement dans le son du film, synthétique, comme « venu d'ailleurs », pour reprendre l'expression de Rudolf Pfenninger<sup>80</sup>, l'un des inventeurs du son optique synthétique (que nous avons évoqué à la section précédente). Les frères Whitney décrivent aux-mêmes ces cinq films dans leurs notes techniques *Notes on the « five abstract film exercises »*<sup>81</sup>. Le premier film (*First sound film*, 1943), ainsi que les deux suivants (*Fragments*, été 1944) emploient la technique des papiers découpés à encoches. Les deuxième et troisième sont restés inachevés, les deux frères ayant décidé de plutôt se concentrer sur les deux suivants.

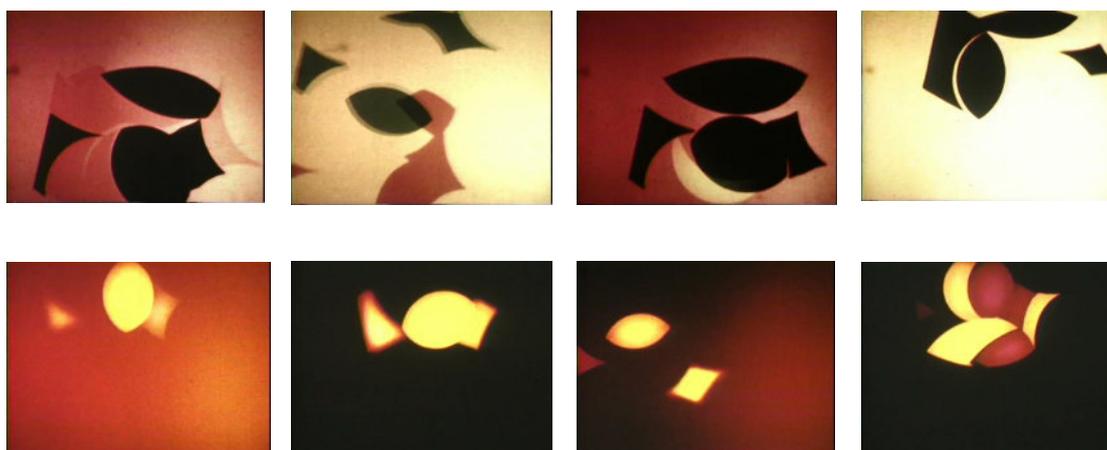


Fig. 13 : John et James Whitney, *Five Abstract Films Exercises* (1943-44)  
© The John Whitney Collection, DVD, Pyramid, 2009

Le quatrième film (*Fourth film*, été 1944) est divisé en quatre sections. La première partie est un jeu sur la profondeur, chaque son est lié à une forme en mouvement qui s'« approche » : plus la forme est proche, plus le son est entendu tôt. Un intervalle tonal est mis en relation à l'espace de l'écran et à l'échelle des formes, les mouvements d'aller-retour des formes vont créer une sorte de glissando hypnotique. Dans la deuxième partie, quatre « sujets » vont être développés dans une alternance de

80 Thomas Y. Levin, « Des sons venus de nulle part. Rudolf Pfenninger et l'archéologie du son synthétique », in Sophie Duplaix, Marcella Lista (dir.), *op. cit.*, p. 51.

81 John et James Whitney, « Notes on the "Five Abstract Film exercises" », in Robert Russett, Cecile Starr (dir.), *Experimental Animation. Origins of a New Art*, Cambridge, Da Capo Press, 1976, p. 173.

contractions et d'expansions mêlées à des changements de rythme. Les sons et les formes graphiques sont joués en totale synchronisation. La troisième partie, qui illustre le dessein principal du film, explore le principe du canon, en lançant toutes les cinq secondes des séquences visuelles et sonores complexes (des « masses » selon les termes de Whitney) qui se simplifient mais accélèrent leur mouvement progressivement. La dernière est dévolue à la répétition et à l'extension du matériel des deux premières parties, avec une inversion menant à sa conclusion.

Le cinquième et dernier film reprend le principe du canon avec un thème de base dont les variations seront la structure même du film. Le son est joué en contrepoint du visuel, mais en suivant aussi un canon. Les dernières séquences jouent de nouveau sur la profondeur de l'écran et les échelles de formes. Même si ce travail était collaboratif, il semblerait que John Whitney ait réalisé le premier et le dernier film seul, tandis que James Whitney aurait réalisé les trois autres. L'ensemble a été récompensé pour sa singularité sonore. William Moritz note à ce propos :

[...] beaucoup de personnes qui furent témoins des premières projections de ces films [...] m'ont dit combien les spectateurs furent bouleversés par la musique « électro-sonique », « supraterrrestre » et les lumineuses images « néon » qui, d'une manière ou d'une autre, semblaient être tombées dans notre sphère temporelle en venant du futur de la science-fiction<sup>82</sup>.

### **Complémentarité audiovisuelle**

On constate à la vue de ces films et à la lecture des notes techniques de leur créateur que les concepts issus des champs musicaux et visuels sont de véritables sources d'inspiration. Le canon par exemple, lié au champ lexical musical, est repris pour mettre en scène un mouvement visuel. Il y a ainsi un va-et-vient conceptuel et structurel entre le son, la forme et le mouvement, qui vient unifier et équilibrer l'ensemble. C'est d'ailleurs l'idée que va continuer à développer Whitney toute sa vie : le son et l'image sont tour à tour en synchronisation ou en contrepoint, selon les différentes phases de la composition, et ils sont « complémentaires » (l'une des composantes ne vient pas

<sup>82</sup> John Whitney, « Musique audio-visuelle : musique couleur — cinéma abstrait », trad. Alain Alcide-Sudre, in Yann Beauvais (dir.), *Musique Film, op. cit.*, p. 26.

prendre le pas l'une sur l'autre, comme le serait un tressage de deux différents éléments) :

Une approche plus fructueuse, moins mécaniste, est possible aujourd'hui parce que de réelles possibilités créatrices émergent quand la structure de l'image dicte ou « inspire » la structure sonore et vice-versa, ou quand leur conception est simultanée. Ceci évidemment, est mieux réalisé quand les deux parties ont des origines créatrices communes. Ce potentiel créateur unifié est aujourd'hui disponible, même pour des artistes, puisque l'écriture de la piste sonore ou les dispositifs des synthétiseurs, maniables par une personne, apparaissent à l'horizon en même temps que des techniques d'animation simplifiées<sup>83</sup>.

Comparé aux idées du début du siècle liées aux correspondances universelles (à savoir qu'à une couleur correspondrait une note), les frères Whitney sont déjà bien en avance car ils définissent un mode de composition au sein duquel l'artiste définit lui-même en amont les correspondances qu'il veut mettre en jeu (comme on l'observe dans le quatrième film avec la mise en relation de la profondeur et de l'intensité du son). Dans ces cinq films, une phase d'« amorce » va servir à mettre en scène les modes d'accroche entre les différentes composantes du film pour être ensuite développées plus en profondeur, une sorte de lexique ou de légende à la manière musicale, dans laquelle un thème est présenté puis développé. Cela sous-entend aussi qu'il n'existe pas de correspondances « universelles ». Il est en revanche possible de construire des entités totalement abstraites, mais lisibles, compréhensibles. Marcella Lista précise que dans ces films :

[...] un registre jusque-là inédit de coïncidences s'observe entre les perceptions auditives et les perceptions visuelles : effets de déplacements sur le plan et dans la profondeur, variation des dimensions, superpositions et rapports optiques de couleurs<sup>84</sup>.

---

<sup>83</sup> *Ibid.*, p. 29.

<sup>84</sup> Marcella Lista, *Empreintes sonores et métaphores tactiles* », in Sophie Duplaix, Marcella Lista (dir.), *op. cit.*, p. 73.

### **Du cinématographique au numérique**

Suite à ces expérimentations, John Whitney va poursuivre seul<sup>85</sup> ses travaux. Pour nous il s'agit d'un artiste très important, puisqu'il va traverser les périodes de l'optique, de l'analogique et du numérique, en étant chaque fois pionnier. Durant toute sa carrière, Whitney va travailler sur des systèmes qu'il aura conçus lui-même ou sur des outils dont les potentialités artistiques n'étaient pas encore explorées. Les premiers films qu'il réalise avec son frère James en sont l'illustration. Il va notamment travailler sur un ordinateur analogique qu'il a créé, puis sera l'un des premiers artistes en résidence à exploiter les potentialités de l'ordinateur. Dans les années 1980, il publie une série de textes décrivant ses recherches dans son ouvrage *Digital Harmony*<sup>86</sup>. Puis il consacre le reste de sa carrière à développer un logiciel de composition sonore et visuelle s'appuyant sur ses propres théories traitant de l'image graphique animée et du son. Nous reparlerons de tous ces travaux au fil des chapitres.

---

85 Il obtient à la fin des années 40 une bourse du Guggenheim et développe pendant deux ans des techniques d'animation en temps réel, à base de papiers découpés et de liquides plus ou moins épais sur lequel il peut dessiner des formes. Il s'intéresse au dessin sur pellicule, comme l'avait fait le cinéaste canadien Norman McLaren. Côté son, il travaille essentiellement avec des musiques préexistantes de Jazz, sur lesquelles il improvise ou interprète des motifs animés en temps réel, qu'il peut ensuite monter ou bien superposer sur la pellicule. Ces techniques plus spontanées lui permettent de réaliser un film complet de trois minutes en un après-midi. Il en réalise une douzaine, qui seront notamment diffusés plus tard à l'exposition universelle et internationale de Bruxelles en 1958, parmi eux *Celery Stalks at Midnight* (1952) ou encore *Dizzy Gillespie Hothouse1* (1952). À la suite de ces expérimentations il cherche ensuite à mettre en place avec l'Université de Californie une sorte d'ensemble constitué d'une demi-douzaine de caméras vidéos et d'une douzaine d'agissants qui expérimenteraient en temps réel les techniques développées dans ses films Jazz, mais sa proposition pourtant innovante ne fût pas acceptée.

86 John Whitney, *Digital Harmony. On the Complementarity of Music and Visual Art*, New Hampshire, McGraw-Hill, Peterborough, 1980.

## 1.1.5 Sons et images des ondes. Physique et mathématiques de l'art

### *Les affinités électroniques*

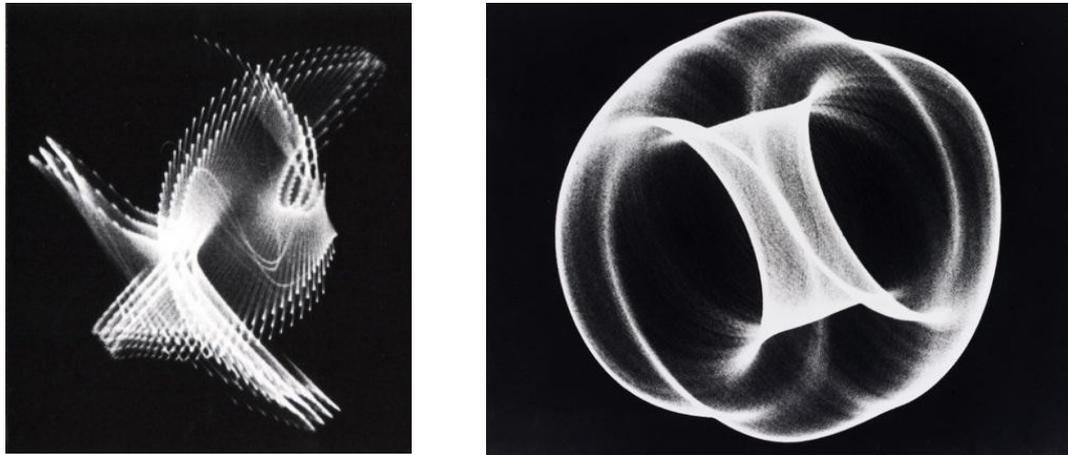


Fig. 14 : Ben F. Laposky, *Oscillons*, circa 1953  
© Centre Pompidou

Après la génération de sons optiques synthétiques, la manipulation de signaux grâce à l'oscilloscope constitue une nouvelle avancée, préfiguratrice du numérique. Bernard Caillaud, dans son ouvrage *La création numérique visuelle*, cite Frank Popper à ce propos :

Le Computer Art — art par ordinateur, ou art informatique — est sans doute né aux États-Unis en 1952, lorsque Ben F. Laposky utilisa une calculatrice analogique et un oscillographe cathodique pour réaliser ses *Electronic Abstractions*. Peu après, en 1956, il réussit à produire une image électronique en couleurs<sup>87</sup>.

Laposky, lorsqu'il décrit ces expérimentations avec l'oscilloscope et d'autres systèmes électroniques, parle de l'idée d'un nouvel art abstrait, basé sur la conception de formes à caractère mathématique. Il y voit aussi un parallèle, voire une potentialité de complémentarité, avec la musique électronique.

Les abstractions, comme nous l'avons montré, sont créées par des ondes électriques, tout comme la musique se compose d'ondes sonores. Les motifs sont abstraits et mathématiques, tout comme la musique est, pour une large part, abstraite et mathématique. Ensuite, l'électronique permet une autre associa-

<sup>87</sup> Bernard Caillaud, *op. cit.*, p. 3.

tion, en ce sens que la musique peut être jouée sur des orgues électroniques ou sur le Theremin, et qu'elle peut même être synthétisée par des machines électroniques<sup>88</sup>.



Fig. 15 : Mary Ellen Bute (circa 1953)

© Paul Young, *Le Cinéma expérimental*, Paris, Taschen, 2009

Les cinéastes d'animation Hy Hirsh, Mary Ellen Bute avec sa série de films « Seeing sound » réalisés au début des années 1950 et Norman McLaren avec *Around is Around* (1951) ont, quant à eux, intégré ce type d'oscillogramme dans certains de leurs films, en filmant directement l'écran de l'oscilloscope. Des pionniers du computer art comme Herbert W. Franke avec ses œuvres de « graphic music » vont aussi explorer ces principes.

La plupart de ces nouvelles formes visuelles inspirées par les mathématiques, sont basées sur ces figures célèbres observées par Jules Antoine Lissajous autour de 1857 et appelées « Courbes de Lissajous ». C'est en partie grâce à des expérimentations avec les signaux d'oscilloscope qu'a émergé cette esthétique visuelle liée à des formes cycloïdales et ondulatoires. Celle-ci va connaître ensuite un nouvel essor avec l'informatique. Il en est de même pour la peinture avec l'art optique et cinétique, qui s'inspire fortement de cette esthétique.

Plus récemment l'artiste audiovisuel Robin Fox, grâce à un système hybride analogique-numérique a expérimenté de nouvelles possibilités de synthèse de ces figures

<sup>88</sup> Ben F. Laposky, « Electronic Abstracts. Art for the Space Age » [Abstractions électroniques. L'art de l'ère spatiale], trad. Jean-François Allain, in Sophie Duplaix, Marcella Lista (dir.), *op. cit.*, p. 21.

dans sa série de vidéos graphiques et sonores *Backscatter* (2004)<sup>89</sup>. Citons aussi Ryoji Ikeda et Carsten Nicolai qui y ont consacré une sorte d'index explorant une multitude de combinaisons de ces figures comme visualisation de signaux sonores dans le cadre de leur projet commun *cyclo* (2000).

### ***Ordinateurs analogiques et imagerie électronique***

À cette période, John Whitney innove encore. À la fin des années 1950, réalisant que ce qui l'intéresse le plus réside moins dans la réalisation que dans le processus de design du contenu qui sera filmé, il commence à fabriquer de nouveaux appareils, les ordinateurs analogiques mécaniques, très complexes à prendre en main. Il va travailler sur ce projet pendant dix ans<sup>90</sup>, faisant évoluer ces appareils afin de créer son propre système d'animation, avec lequel il va produire une série d'œuvres réunies sous le titre *Catalog* (1961)<sup>91</sup>.



Fig. 16 : John Whitney, *Catalog* (1961)  
© Rhizome.org

Il y a chez tous les artistes que nous évoquons la volonté de fabriquer ou « générer » leurs propres images et leurs propres sons, sans se servir d'enregistrements du réel. C'est pour cela qu'il existe une véritable continuité artistique entre tous ces travaux, qui tendent à s'extraire du monde réel, à s'émanciper du modèle.

89 Jan Thoben « Technical Sound-Image Transformations », in Dieter Daniels, Sandra Naumann (dir.), *Audiovisuology Compendium : An Interdisciplinary Survey of Audiovisual Culture*, Köln, Walther König, coll. « See This Sound », 2010, p. 425.

90 Son procédé devenant populaire, il crée en 1960 sa propre compagnie Motion Graphics, Inc. et va réaliser une série de films publicitaires et commerciaux, dont une des parties graphiques (la spirale) du générique de *Vertigo* (1958) d'Hitchcock en collaboration avec le réalisateur Saul Bass. Ces animations sont remarquables, mais l'aspect sonore étant laissé de côté par Whitney pendant cette période, nous ne la détaillerons pas ici. Notons que le frère de John, James Whitney, continue quant à lui à réaliser patiemment de son côté des films optiques nécessitant un travail très précis et très long, dont son plus connu *Lapis* (1966).

91 Une version est disponible en ligne sur Rhizome.org, <https://rhizome.org/editorial/2008/nov/21/catalog-1961-john-whitney/>, consulté le 30/01/23.

## 1.2 *Le son des premiers films programmés*

### 1.2.1 L'ordinateur avant l'ordinateur : une machine analytique

L'ordinateur, tel que nous le connaissons, est le fruit d'une longue évolution des machines à calculer vers ce qui sera nommé, sur la base de l'article « On Computable Numbers »<sup>92</sup> d'Alan Turing en 1936, la machine de Turing universelle. Il y décrit théoriquement un appareil de calcul qui utiliserait la lecture et l'écriture de données sur un ruban sans fin.

L'un des premiers ancêtres de cette machine universelle fut la machine analytique, inachevée, de Charles Babbage en 1833<sup>93</sup>. Celle-ci était un calculateur à base de cartes perforées contenant un programme, ainsi que des données à traiter. De ce principe de stockage sur un support amovible et interchangeable résulte une forme primaire de programme. Ada Lovelace, qui collabore avec Babbage, formalise déjà des séries d'instructions qui pourraient rendre automatique un processus, se rapprochant de l'ébauche d'un premier langage de programmation. L'idée originelle de l'utilisation de la carte perforée de Babbage proviendrait du métier à tisser inventé par Joseph Marie Jacquard vers 1800. C'est ce qui fait dire à Lev Manovich qu'une machine à cartes perforées « synthétisait [...] déjà des images avant de servir au traitement des nombres<sup>94</sup> ». Nous ajouterons à cela que les cartes perforées étaient déjà aussi utilisées dans la musique. Elles servaient de base pour des instruments mécaniques automatisés comme le célèbre orgue de Barbarie et beaucoup d'autres, qui lui étaient antérieurs. Parmi eux, il est notable que le comptonium de Winkel en 1821<sup>95</sup> intègre déjà une part d'aléatoire, grâce au jeu de deux cylindres reposant sur des mécanismes d'horlogerie. Il ne s'agit pas de réel hasard, mais d'une possibilité étendue de variations combinatoires.

En 1890, Herman Hollerith invente une machine fondée sur l'utilisation étendue de ces cartes perforées, appelée tabulateur. Elle servait initialement pour le recensement aux États-Unis, où elle connaît un vif succès. Il crée une société qui se développe et fu-

92 Alan Turing, « On computable numbers, with an application to the Entscheidungsproblem », in *Proceedings of the London Mathematical Society*, vol. 42, n° 2, Londres, 1936.

93 Lev Manovich, *Le langage des nouveaux médias*, Dijon, Les Presses du réel, 2010, p. 93.

94 *Ibid.*, p. 91.

95 Moreno Andreatta, « Coder le monde de la musique : une visite guidée à travers la timeline musique », in Frédéric Migayrou (dir.), *Coder le monde*, Orléans, HYG / Centre Pompidou, coll. « Script », 2018, p. 147.

sionne avec d'autres pour devenir, en 1924, IBM<sup>96</sup>. Cette compagnie s'impose comme un acteur important de l'émergence de l'art par ordinateur dans les années 1960.

Le premier ordinateur programmable est baptisé le Z3, inventé en 1941 par un ingénieur allemand, Konrad Zuse. Fait notable et symbolique, il emploie comme ruban de stockage des chutes de films 35 mm qui se voyaient perforés pour devenir support de données<sup>97</sup>. À la suite du Z3 de Zuse, plusieurs modèles d'ordinateur, de plus en plus puissants, vont faire leur apparition. L'un des plus connus est l'ENIAC<sup>98</sup>, en 1947. La plupart étaient destinés à une exploitation militaire. À la sortie de la guerre, de grands laboratoires industriels et quelques grandes universités font l'acquisition de machines informatiques. À partir de ce moment précis, certains artistes vont pouvoir investir artistiquement ce matériel, souvent en collaboration avec des ingénieurs.

### 1.2.2 Émergence de la musique informatique et de l'art par ordinateur

Si l'on croise l'Histoire de la musique et celle du visuel par ordinateur, on s'aperçoit que le son devient numérique avant l'image (fixe), devant son apparition de presque dix ans.

#### *La musique et l'informatique*

L'éclosion de l'informatique musicale débute en 1955<sup>99</sup> avec le premier synthétiseur programmable (appelé le RCA Mark II Sound Synthesizer). En 1957, la partition d'une composition musicale pour quatuor à cordes, *Illiad Suite for String Quartet* (Lejaren Hiller et Leonard Isaacson), est conçue à partir d'un ordinateur.

---

96 Acronyme de International Business Machines.

97 Lev Manovich, *op. cit.*, p. 95.

98 Electronic Numerical Integrator And Computer. Il s'agit du premier calculateur totalement électronique.

99 D'autres expériences préexistent, mais celles-ci restent à l'état de tests : à Melbourne, puis à Manchester.



Fig. 17 : IBM 704 Electronic Data Processing System (1957)  
© IBM, Computer History Museum

La même année, l'ingénieur Max Mathews<sup>100</sup> achève le premier langage de musique informatique aux laboratoires Bell (les Bell Labs), *MUSIC I*. Il s'agit d'un programme en assembleur, développé avec un ordinateur IBM 704, qui donnera naissance à la série *MUSIC-N*<sup>101</sup> de logiciels (dotés, à partir de 1968, d'une interface graphique). Mathews fait venir aux laboratoires Bell des compositeurs afin qu'ils puissent expérimenter avec cet outil informatique inédit. Le premier morceau composé, d'une durée de 17 secondes, est *The Silver Scale*, du compositeur Newton Gutmann, puis suit la même année *Pitch Variations*. En 1960, la compilation *Music for Mathematics* voit le jour. Elle regroupe les pièces pionnières de cette période. Au milieu des années 1960, le français Jean-Claude Risset intègre le studio à son tour. Il y fait naître une démarche inédite croisant arts, sciences et psychologie cognitive. Alors que se développe une branche de la synthèse sonore qui cherche à imiter les sons existants des instruments, Risset prône la composition de nouvelles natures et timbres de sons, pour faire émerger de nouvelles pratiques. Dans l'un de ses textes, Risset évoque Varèse, qui avait le rêve de se départir de la lutherie, et du tempérament égal<sup>102</sup> qu'il considérait comme un « fil à couper l'octave » :

100 Le « Max » figurant dans le nom du logiciel de programmation graphique *Max/MSP/Jitter*, développé par Miller Puckette dans les années 1980, est un hommage à Max Mathews.

101 Le N définissant la version. La dernière version programmée par Mathews est *MUSIC V* (1968). La dernière version de la série des *MUSIC-N*, *MUSIC 11* (1973), a été développée par Barry Vercoe. Celui-ci a décliné cette version en un nouveau logiciel, *Csound*, écrit en langage C, au MIT en 1986. *Csound* est encore très utilisé aujourd'hui.

102 C'est-à-dire la gamme tempérée, qui divise l'octave musicale en douze intervalles égaux.

Mieux que les autres ressources extramusicales, l'ordinateur permet d'envisager la réalisation de ce rêve : il est capable de fournir un matériau d'une immense variété et d'une extrême malléabilité. Et plus peut-être que la nouveauté des sons qu'il peut produire — la nouveauté d'un son est éphémère — c'est la finesse du contrôle musical qu'il permet d'exercer sur ces sons qui en fait un instrument unique, à même d'ouvrir de nouvelles voies à la création musicale<sup>103</sup>.

En travaillant notamment sur les sons paradoxaux (les illusions sonores), il fait évoluer la fameuse gamme de Shepard, illusion d'un son montant sans cesse à l'infini. Il met aussi au point un premier solfège de sons synthétiques<sup>104</sup>, dans une démarche pédagogique. Au même moment, le compositeur Iannis Xenakis publie son ouvrage *Musiques formelles*<sup>105</sup> (1963), dans lequel il documente sa méthodologie de composition avec des algorithmes et des probabilités.

### ***L'art visuel par ordinateur***

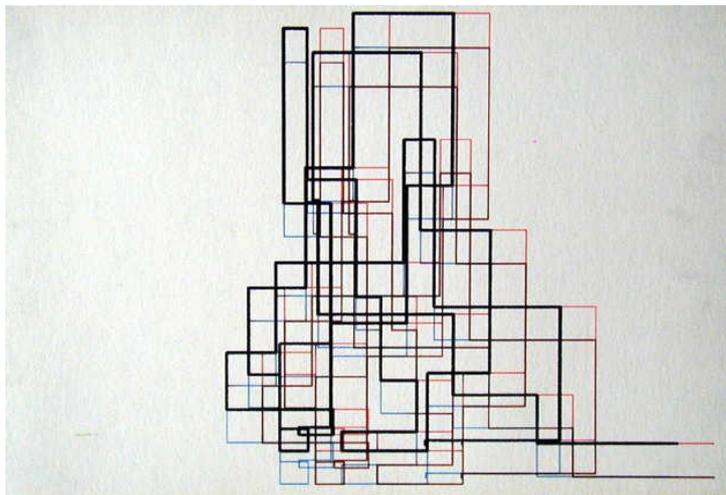


Fig. 18 : Frieder Nake, 105/130 (1965)  
© compArt, Center of excellence digital art

La plupart des ordinateurs disponibles dans les années 1960 sont des *Mainframe*<sup>106</sup>, c'est-à-dire d'énormes unités de calcul coûteuses et imposantes. Jasia Reichard, qui

103 Jean-Claude Risset, *op. cit.*, p. 23.

104 Jean-Claude Risset, *An Introductory Catalogue Of Computer Synthesized Sounds*, 1969.

105 Iannis Xenakis, *Musiques formelles*, Paris, *Revue Musicale* n° 253-254, Stock, 1963.

106 L'ouvrage *Mainframe Experimentalism. Early Computing and the Foundations of the Digital Arts* s'articule autour de cette terminologie du *Mainframe* (Hannah B. Higgins, Douglas Kahn [dir.], *Mainframe Experimentalism: Early Computing and the Foundations of the Digital Arts*, Berkeley, University of California Press, 2012).

fut la commissaire de l'exposition *Cybernetic Serenpidity*<sup>107</sup> à l'ICA en 1968, souligne que ces machines n'étaient accessibles que par le biais de centres de recherche de grands industriels, d'universités avec un centre informatique, ou bien encore par les sociétés d'informatique<sup>108</sup>. Les artistes devaient donc solliciter des résidences pour avoir accès à ce matériel, souvent assistés d'ingénieurs. Jasia Reichard prend soin de rappeler que l'art par ordinateur n'était pas seulement visuel, mais concernait aussi la musique, ainsi que les formes textuelles et poétiques (ce qui est bien représenté dans l'exposition *Cybernetic Serenpidity*).

En 1963, après la musique, l'éventail informatique s'étend à l'image. Plusieurs pôles de recherche et de création émergent dans différents pays (le Royaume-Uni, l'Allemagne, l'Italie, l'Autriche, les USA, le Japon, le Canada ou encore l'Amérique du Sud). En Allemagne, les premières images par ordinateur sont réalisées par Frieder Nake et Georg Nees<sup>109</sup>. De manière concomitante aux États-Unis, se forme aux laboratoires Bell une équipe (constituée entre autres de Ken Knowlton, Béla Julesz et Michael Noll) qui va travailler sur les premiers langages de génération graphique. Au sein de ces pôles d'émergence, des collectifs de recherche et de création entre arts et technologies<sup>110</sup> voient le jour (comme la Computer Arts Society au Royaume-Uni en 1968, le Computer Technique Group — CTG — au Japon en 1969, ou bien le Computer Graphics Research Group — CGRG — aux États-Unis). Fait étonnant, alors que la musique informatique fût relativement bien accueillie, l'art visuel par ordinateur reçut un accueil très mitigé de la part de la critique<sup>111</sup>. Pourtant, une nouvelle forme de création graphique « générative » et « programmable » était née. Manfred Mohr, l'un de ces pionniers, s'en explique à propos de son propre travail :

107 Exposition *Cybernetic Serenpidity*, Institute of Contemporary Arts, Londres, du 2 août au 20 octobre 1968. Commissaire : Jasia Reichard.

108 Jasia Reichard, *The Computer in Art*, Londres, Studio Vista, 1971, p 8.

109 C'est également cette année-là que Ivan Sutherland invente le programme informatique *Sketchpad*. Il s'agit de la première fois qu'est utilisée une interface, avec un moniteur et un stylo optique. IBM commercialise, un an après, son premier modèle intégrant ces outils externalisés, l'IBM 2250.

110 Sur, en quelque sorte, le modèle de ce qui avait été initié par l'EAT (Experiments in Art Technology), organisation créée en 1966 par des ingénieurs (Billy Klüver et Fred Waldhauer) et des artistes (Robert Rauschenberg et Robert Whitman).

111 Edmond Couchot et Robert Hillaire, *L'art numérique. Comment la technologie vient au monde de l'art*, Paris, Flammarion, coll. « champs/arts », p. 22.

L'analyse d'une idée visuelle fait apparaître des structures internes qui peuvent être considérées comme les matériaux élémentaires de cette structure. Ces éléments permettent la construction de systèmes dans lesquels se développe une pensée logique, c'est-à-dire une conception programmable de l'idée originelle. Les dessins générés par de tels processus logiques, dont le déploiement n'est que partie d'une structure globale, sont appelés dessins génératifs. Ce qui caractérise ces systèmes génératifs, c'est la multiplicité des dessins uniques qu'ils engendrent. La construction syntactique<sup>112</sup> de l'œuvre d'art devient ainsi le principe d'une compréhension visuelle et intellectuelle<sup>113</sup>.

Les théoriciens et chercheurs Edmond Couchot et Norbert Hillaire parlent, quant à eux, de l'ordinateur comme une « machine de *feed-back* qui rétroagit sur l'imagination technique de l'artiste et ouvre à celle-ci des possibilités insoupçonnées au départ par l'artiste lui-même<sup>114</sup> ».

### ***De l'image fixe à l'image animée : les premiers films par ordinateur***

La première décennie de l'art visuel par ordinateur est celle d'une exploration de l'abstraction graphique par la génération programmée de formes géométriques primaires, en noir et blanc. La maîtrise d'un langage proche de la machine<sup>115</sup>, tel que le FORTRAN<sup>116</sup>, est nécessaire. Le FORTRAN a été utilisé à la fois par des musiciens et par des artistes visuels, de différentes manières. Comme nous l'avons dit, les ordinateurs sont coûteux et imposants, et ne permettent pas un travail en temps réel. Il faut programmer, fabriquer la carte perforée contenant le programme, traiter ce dernier et imprimer le résultat visuel, fixe, à l'aide, par exemple, de *plotters* [traceurs].

Les images générées par ordinateur sont alors fixes. C'est en travaillant image par image, ou bien en exposant des suites d'images sur microfilm pour les passer en film 16 mm, que les premiers films « programmés » vont faire leur apparition. À ce titre

---

112 Synonyme de « syntaxique ».

113 Manfred Mohr, « Artist's Statement », in Manfred Mohr, *CUBIC LIMIT*, Paris, Galerie Pierre Weiller, 1975.

114 Edmond Couchot, Norbert Hillaire (dir.), *op. cit.*, p. 20.

115 Un enregistreur de microfilms pouvant être contrôlé par ordinateur.

116 Abréviation de FORMula TRANslator. Le langage a été créé dans les années 1950.

l'une des machines fondamentales fût le Stromberg-Carlson S-C 4020<sup>117</sup> : elle permit ce passage du numérique au film.

On cite souvent *A Two Gyro Gravity Gradient Altitude Control System* (EUA, 1962-1963) de Edward Zajec comme le premier film programmé. Il est toutefois plus une simulation scientifique qu'une véritable pièce artistique. C'est aussi le cas de *Simulated Basilar Membrane Motion*, une animation de A. Michael Noll, qui montre le fonctionnement de la membrane de l'oreille interne<sup>118</sup>. La plupart de ces animations sont encore muettes.

Bien évidemment, la question du son et de l'image va se poser plus fortement avec ces premières images animées générées par ordinateur, la temporalité induite par l'image offrant un espace ouvert au sonore.

Nous avons étudié la documentation des premières expositions de l'époque, ainsi que les archives des groupes de recherche ou d'artistes individuels, à la recherche de pièces informatiques dans lesquelles l'image et le son dialoguait. Nous en donnons, dans les sections qui suivent, les exemples les plus représentatifs.

---

117 Cette machine est le cœur de l'ouvrage de Zabet Patterson, *Peripheral Vision : Bell Labs, the S-C 4020, and the Origins of Computer Art*, MIT Press, 2015.

118 Philippe Bootz, « Quelle est l'histoire de l'animation numérique ? », Leonardo/Olats, décembre 2006, [http://archive.olats.org/livresetudes/basiques/animationnumerique/3\\_basiquesAN.php](http://archive.olats.org/livresetudes/basiques/animationnumerique/3_basiquesAN.php), consulté le 30/01/23.

### 1.2.3 *Cibernetik 5.3*, un système audio-graphique cellulaire

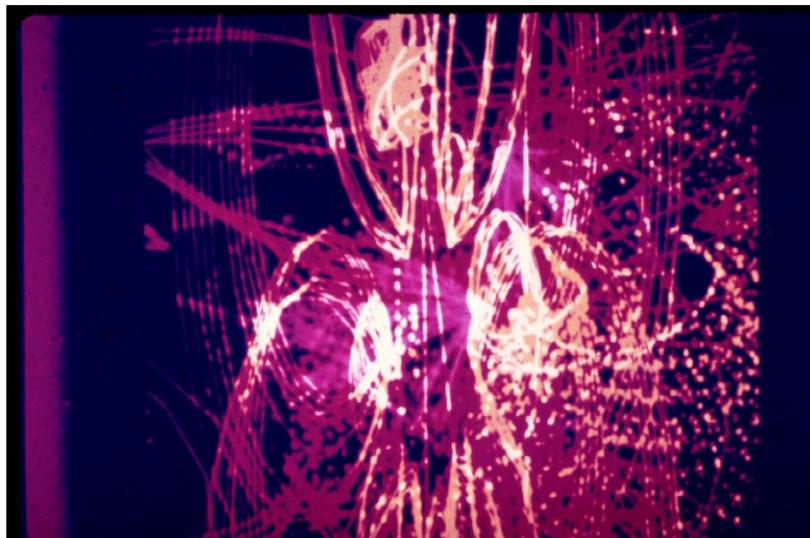


Fig. 19 : John Stehura, *Cibernetik 5.3* (1960-1965)  
© John Stehura, cyberanimation.tripod.com

*Cibernetik 5.3* (1960-1965), de John Stehura, est l'un des tout premiers films artistiques réalisés grâce à un ordinateur<sup>119</sup>. Il s'agit d'une pièce d'environ 8 minutes, dans laquelle des groupes de points et des lignes se déploient, se superposent, se regroupent, disparaissent, à la manière de mouvements cellulaires<sup>120</sup>. La couleur y est très prégnante. Peu à peu viennent se mêler des images figuratives, notamment des photographies de visages. Tout n'est donc pas abstrait mais les deux tiers du film sont constitués de compositions graphiques génératives animées qui s'appuient sur des algorithmes de déplacement et de transformation.

Pour cette création, Stehura utilise la pièce *Quatermass* de Tod Dockstader, sortie officiellement un peu plus tard en 1966. Il s'agit d'une œuvre de 45 minutes de musique concrète — c'est-à-dire composée d'échantillons de sons enregistrés avec un micro-

---

119 Le titre est souvent mal orthographié, il s'agit bien de *Cibernetik 5.3*. La date du film qui figure sur les programmes ou circule sur le Web est la plupart du temps 1965-69. Nous avons pu échanger par email avec Cindy Keefer, directrice du CVM (Center for Visual Music) à Los Angeles. Cette structure, à but non lucratif, conserve, préserve et distribue des films d'animation expérimentale associés à la musique visuelle, comme ceux d'Oskar Fischinger ou bien de Jordan Belson. Ayant travaillé avec le réalisateur pour une restauration, elle nous a informé que le film de John Stehura datait bel et bien de 1960-65. Elle a également pu nous mettre à disposition une version digitale de cette version restaurée, qui offre une qualité superbe qui met en valeur les mouvements cellulaires qui composent en partie ce film précurseur.  
120 Le court extrait d'une version restaurée est disponible en ligne sur le Facebook du Center for Visual Music, <https://www.facebook.com/watch/?v=2443493175964260>, consulté le 30/01/23.

phone, mais dont l'attaque (le tout début du son) a été tronquée<sup>121</sup>. Stehura n'en exploite qu'un fragment, intitulé *Tango*. Même si la musique est ajoutée dans un second temps, l'ensemble image-son fonctionne car la nature abstraite des deux composantes appelle notre perception à créer des liens entre elles.

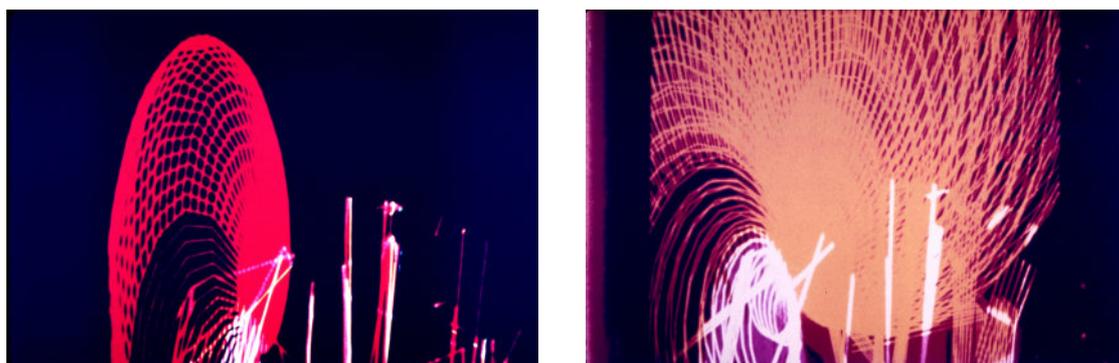


Fig. 20 : John Stehura, *Cybernetik 5.3* (1960-1965)  
© John Stehura, cyberanimation.tripod.com

Dans la version restaurée que nous avons pu visionner et écouter, la programmation générée des mouvements, vraiment novatrice, annonce l'utilisation d'algorithmes de modélisation physique de mouvements pour des formes graphiques. La bande sonore, sorte de battement de cœur sourd et crépitant, se fait entendre avant que l'image ne s'anime. Des grappes de lignes vert clair apparaissent sur un fond vert foncé. Certaines grappes sont constituées de lignes pointillées, d'autres ont des échelles différentes, ce qui crée une profondeur. Au bout d'une trentaine de secondes, le son (une sorte de vibration de lame) devient plus présent. Puis la couleur du fond change, celles des lignes aussi. D'autres groupes de formes abstraites, plus courbes, font leur apparition. Chaque grappe semble évoluer de manière autonome, sans vraiment se préoccuper des autres, les densités graphiques varient avec l'intensité du son. Au bout de 3 minutes environ, des grondements surviennent, accompagnés d'explosions graphiques. Des spirales de lignes s'agrandissent pour prendre possession de l'écran. Des jeux sur le continu et le discontinu s'opèrent, des mouvements particuliers et tourbillonnaires s'entremêlent dans des clignotements d'arrière-plan. À 5 minutes, un son proche de

<sup>121</sup> Cette méthode, inventée par Pierre Schaeffer au début des années 1950, rend le son original non reconnaissable. Il parle à ce propos d'objet sonore, un son qui n'est écouté que pour ses qualités morphologiques, et non pas relativement à sa source de fabrication.

celui d'une alarme retentit. L'image figurative (une sphère dans laquelle des visages sont déformés) surgit, puis quelques minutes plus tard disparaît. Les formes graphiques réapparaissent, une dernière fois avant un passage au blanc et à la lumière qui clôture le film.

À son arrivée à l'UCLA (l'université de Californie), John Stehura, déjà passionné de cinéma, s'intéresse aux recherches naissantes de Watson et Crick sur le génome (ADN). Il découvre toute l'étendue des possibilités naissantes offertes par la création d'images grâce à l'ordinateur et cherche à mettre au point un générateur de film d'animation. Il commence alors à programmer sur un IBM 7094<sup>122</sup> (un *mainframe*), en FORTRAN IV. Ce type d'ordinateur, très coûteux n'avait aucune interface graphique. Stehura a donc travaillé de manière algorithmique, sur le modèle d'une partition mais sans en connaître exactement le rendu. Sur la base de ce langage informatique, il a développé sa propre couche de programmation inspirée de modèles cellulaires, cherchant à générer des environnements, sortes d'organismes abstraits à comportements. Pour cela, il imagine un système de plusieurs modules. L'un génère ou détruit des formes à partir d'algorithmes, un autre produit des systèmes qui influent sur ces ensembles de formes (à la manière du vent sur les feuilles), le dernier permet de créer des transformations (perspectives, lumières, etc.). Le tout est contrôlé par une sorte d'intelligence artificielle. Entre 1960 et 1964, il teste différents algorithmes, créant chaque année un nouveau système de génération d'animation. Il travaille seul, surtout la nuit après les cours, car son approche est trop éloignée des méthodologies scientifiques de l'époque.

La stabilisation d'une première version de son système *Cybernetik* (Model 5, Version 3) va lui permettre, au terme de plus d'une dizaine d'heures de calcul, de produire les images du film. Puis, deux semaines environ plus tard, grâce à l'enregistreur digital Stromberg-Carlson SC-4020 situé dans un autre centre, de récupérer des films 16 mm sur lesquels il verra pour la première fois ses suites d'images générées, en noir et blanc.

---

122 Il s'agit du même ordinateur grâce auquel la première synthèse vocale sera diffusée en 1961, dans le morceau *Daisy Bell*. Nous en parlerons dans la troisième partie, dans le chapitre qui traite de la synthèse vocale.

L'aspect sonore a été pensé très en amont dans le système de Stehura : il a prévu une interface (au sens de liaison) entre son logiciel et le logiciel de génération de son *MUSIC IV*<sup>123</sup>. Chaque objet graphique mis en mouvement pouvait potentiellement être couplé à un son de synthèse, sons dont les paramètres sonores (l'attaque, la chute, l'entretien et l'extinction<sup>124</sup>) pouvaient être contrôlés par *MUSIC IV*. Il semble avoir programmé dans le logiciel plus de 25 instruments<sup>125</sup> (c'est-à-dire 25 sons ayant chacun un timbre différent). À partir du programme de base, Stehura a réussi à générer un fragment de la bande sonore synthétique synchronisable avec l'image. Il la décrit comme un ensemble de chants d'étranges oiseaux interplanétaires dans des arbres<sup>126</sup>. Toutefois, pour que le rendu sonore soit plus intéressant, il aurait fallu fournir un travail encore plus important et long, aussi John Stehura décide d'intégrer une bande-son préexistante, qu'il emprunte à Tod Dockstader et qu'il ajoute peu après. Même si le procédé n'a pas été mené à son terme, Stehura a tout de même probablement été le premier à faire fonctionner un système de correspondance informatique entre des paramètres de mouvement et des paramètres sonores.

---

123 Il s'agit de la quatrième version de *MUSIC-N* créé aux Bell Labs par Max Mathews.

124 En musique informatique, ces paramètres sont ceux, fondamentaux, de l'enveloppe ADSR (Attack, Decay, Sustain, Release) d'un son.

125 John Stehura, « Home Page for a Cybernetic Cinema », <https://cyberanimation.tripod.com/historypage.htm>, consulté le 30/01/23.

126 *Ibid.* « which sounded like strange interplanetary birds in trees ». Nous traduisons.

## 1.2.4 Harmonie digitale. Les films par ordinateur de John Whitney

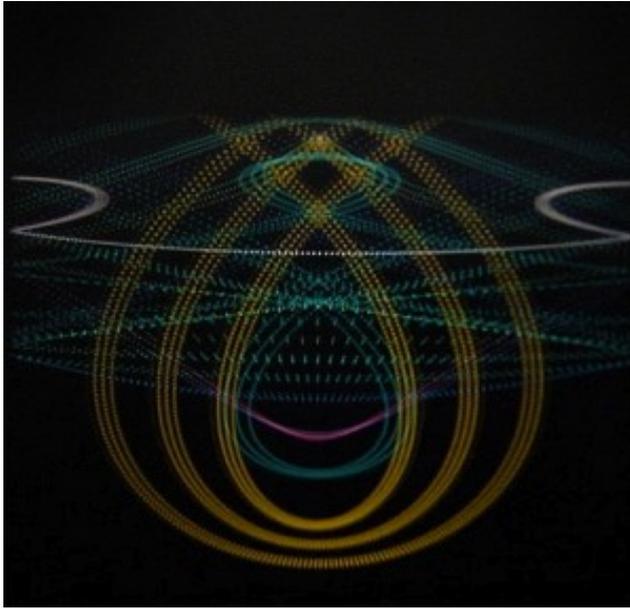


Fig. 21 : John Whitney, *Arabesque* (1975)  
© The John Whitney Collection, DVD, Pyramid, 2009

### « *Experiments in motion graphics* »

En 1966, John Whitney fait une proposition de recherche à IBM. Sa renommée, couplée au soutien du designer Charles Eames, lui permet d'obtenir une bourse et un accès aux studios informatiques. Il est ainsi le premier artiste en résidence chargé d'explorer les possibilités expressives des nouveaux ordinateurs IBM : l'« IBM Model 360 » et la console graphique IBM 2250 (IBM 2250 Graphic display console).



Fig. 22 : Le Dr Jack Citron devant la console graphique IBM 2250  
© The John Whitney Collection, DVD, Pyramid, 2009



Fig. 23 : Changement des valeurs de variables sur l'écran tactile grâce à un stylet  
© The John Whitney Collection, DVD, Pyramid, 2009

Il collabore avec le docteur Jack Citron, d'IBM, et, ensemble, ils vont réaliser des séries d'expérimentations qui donneront naissance aux films les plus importants de Whitney : *Permutations* (1968), la série de trois films *Matrix* (1971-72) et *Arabesque* (1975). D'une certaine manière, le matériel étant prêt à l'emploi, Whitney n'a cette fois pas à inventer de nouvelles machines et, moins préoccupé par la technique, il va connaître une période dense de pure créativité.

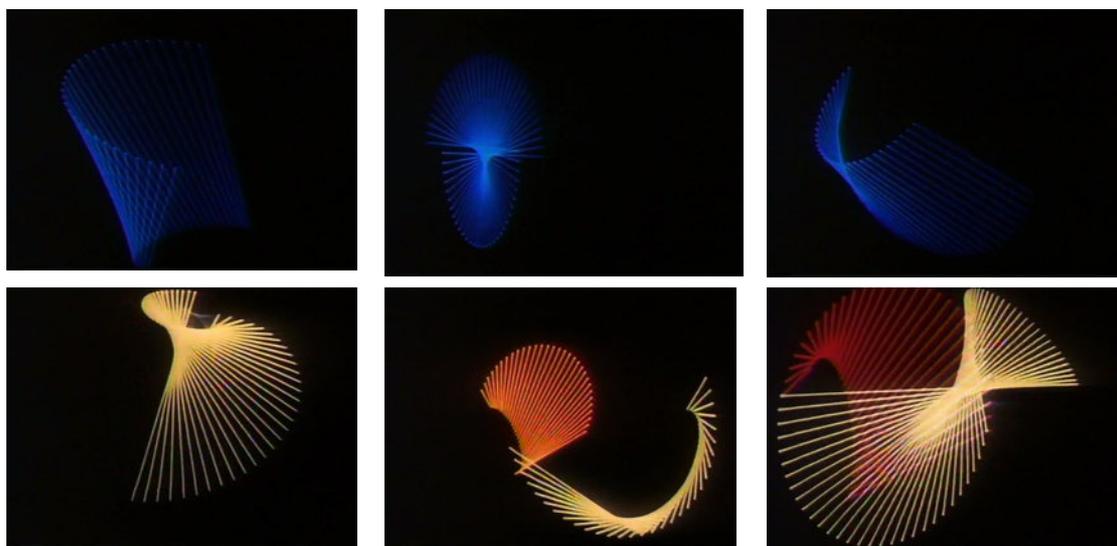


Fig. 24 : John Whitney, *Experiments in motion graphics* (1968)  
© The John Whitney Collection, DVD, Pyramid, 2009

Dans un petit film documentaire de 13 minutes, *Experiments in Motion Graphics* (1968), il commente son processus de travail et ses concepts d'animation des formes graphiques. Les études animées qu'il montre sont parmi les premières jamais réalisées à l'aide d'un ordinateur. Il s'agit de formes géométriques constituées de lignes en rotation se mouvant en permanence tandis que le centre de gravité se déplace. Une sorte de perfection esthétique se dégage de ce ballet de lignes, qui bien que totalement abstrait, provoque une certaine émotion, semblable à celle éprouvée devant une peinture abstraite. Cet art du mouvement pur se manifeste ici par le biais d'une machine générative. Whitney se rapproche du mouvement abstrait, et investit de manière spécifique et personnelle l'espace de l'écran.

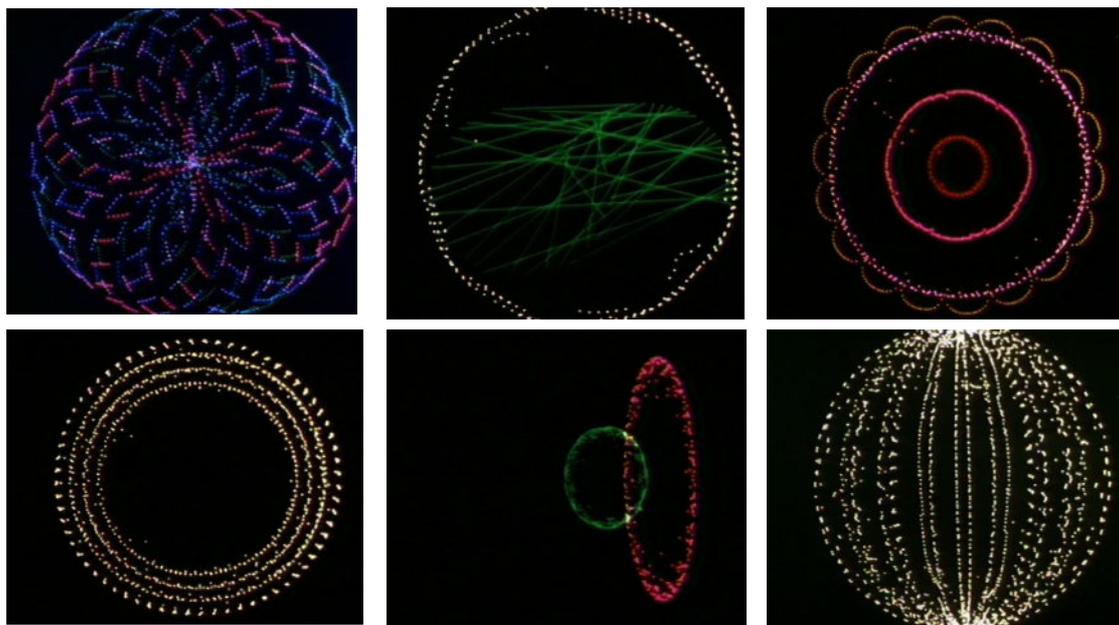


Fig. 25 : John Whitney, *Permutations* (1968)  
 © The John Whitney Collection, DVD, Pyramid, 2009

*Permutations* (1968)<sup>127</sup> est un film de 8 minutes, composé d'une masse de 281 points colorés opérant des mouvements circulaires. Dans ce film, Whitney explore les moyens par lesquels les points en mouvement peuvent produire des effets perceptuels similaires aux modulations dans le champ musical. Ce film est l'un des plus renommés de John Whitney. Il est aussi le premier film entièrement généré par ordinateur. Comme l'explique toutefois Whitney<sup>128</sup>, il ne s'agit pas d'une production en temps réel, chaque image est générée par l'ordinateur une à une, puis transférée depuis l'écran monochrome vers un film 16 mm. Deux secondes de temps de calcul étaient nécessaires pour créer chaque image. La musique du film, rythmée sur des tablas indiens, a quant à elle été ajoutée après le montage du film. Côté programmation informatique, *Permutations* utilise la bibliothèque GRAF (basé sur le langage FORTRAN), développée par le Dr Jack Citron d'IBM<sup>129</sup>.

Comme nous l'avons vu, une grande partie des pionniers du « computer art », tels que Georg Nees, Frieder Nake, Manfred Mohr, or Chuck Csuri, des contemporains de John

127 Une version est disponible en ligne sur Internet Archive, [https://archive.org/details/permutations\\_201608](https://archive.org/details/permutations_201608), consulté le 30/01/23.

128 John Whitney, *A personal search : for the complementary of Music and Visual Art*, 1992, in *The John Whitney Collection*, DVD, Pyramid Media, 2009.

129 Casey Reas, Chandler McWilliams (dir.), *Form+Code in Design, Art, and Architecture*, New York, Princeton Architectural Press, 2010, p. 23.

Whitney, tireront aussi partie du langage FORTRAN. Ces derniers seront toutefois plus intéressés par les possibilités artistiques apportées par la génération programmée d'images fixes et de leur impression sur table traçante, tandis que Whitney se passionne pour les qualités musicales du mouvement. Ceci s'explique bien entendu par son expérience dans le film et l'animation, et lui confère une certaine avance dans ce domaine.

Chaque point dans *Permutations* se déplace à une vitesse différente et se dirige dans une direction indépendante en accord avec les lois naturelles... Leur action produit un phénomène qui est plus ou moins équivalent aux harmoniques musicales. Lorsque les points atteignent certains rapports numériques avec d'autres paramètres de l'équation, cela forme des courbes géométriques élémentaires pouvant rappeler les courbes d'une rosace. Ensuite, ils repartent dans un rapport et apparaissent aléatoirement à nouveau<sup>130</sup>.

*Matrix* (1971-72)<sup>131</sup> est une série de trois films. Le premier, *Matrix I*, et le deuxième, *Matrix II*, sont basés sur les mêmes animations graphiques, mais avec une musique différente. *Matrix III* est constitué de plusieurs sections dans lesquelles une forme élémentaire est dupliquée au gré de variations animées se basant sur des trajectoires elles aussi élémentaires. Par exemple, une des sections nous montre les variations des clones d'un même triangle dont les trajectoires sont des cercles, Whitney ménageant des points de « résonance » durant lesquels l'ensemble de formes atteint des points culminants, que Whitney nomme « nœuds harmoniques<sup>132</sup> ». La musique, ajoutée aussi après le montage, est tirée de l'œuvre du compositeur minimaliste Terry Riley *Rainbow in Curved Air* (1969), écrite pour orgue électronique et bande magnétique.

Le dernier film, *Arabesque* (1975)<sup>133</sup>, est un film de 7 minutes inspiré par les ornements. La musique est de Manoocheher Sadeghi. Le design génératif de l'œuvre est constitué de 360 points distribués autour d'une structure circulaire et qui vont entrer en mouvement selon des règles de calcul très précises. Les formes composées de ces

130 John Whitney, *Digital Harmony*, op. cit., p. 197.

131 Des versions sont disponibles en ligne sur Internet Archive, [https://archive.org/details/matrix\\_201703](https://archive.org/details/matrix_201703), consulté le 30/01/23.

132 *Ibid.*, p. 75.

133 Une version est disponible en ligne sur le site du SIGGRAPH, <https://digitalartarchive.siggraph.org/artwork/john-whitney-arabesque/>, consulté le 30/01/23.

points, au départ très simples, vont se développer, se superposer et l'on va assister à un contrepoint de différents mouvements oscillatoires, circulaires ou translatoires, certains étant imbriqués les uns dans les autres. Les changements de couleur permettent de différencier les niveaux. L'ensemble est un véritable tour de force complexe et hypnotique. Nous ne pouvons pas ici en détailler tous les mouvements, mais John Whitney l'a fait dans son ouvrage *Digital Harmony*<sup>134</sup>, dont nous reproduisons ici quelques extraits. Il réalise ce film avec le programmeur et futur cinéaste Larry Cuba<sup>135</sup>. Les images sont pour la première fois enregistrées à l'aide d'une imprimante électronique, et non plus à l'aide d'une caméra 35 mm.

Dans ces premiers films par ordinateur, la relation entre image et musique est presque intuitive, rythmique, et donc beaucoup moins inventive que la série de films audio-optiques *Five Abstract Film Exercises* des années 1940 réalisés avec son frère<sup>136</sup>. En revanche, le modèle musical et mathématique utilisé pour la mise en mouvement — son harmonie digitale — est travaillé avec une grande précision.

---

134 *Ibid.*, p. 97.

135 Larry Cuba va réaliser 3 films graphiques et sonores par ordinateur : *3/78 (Objects and Transformations)* (1978), *Two Space* (1979), *Calculated Movements* (1985). Tous s'inscrivent dans une esthétique minimaliste, en noir et blanc. Nous évoquerons son travail dans la quatrième partie. Larry Cuba a récemment remasterisé ces 3 films en numérique. Il considère que ce sont ces nouvelles versions qui doivent être visionnées.

136 Nous en avons parlé de cette série au premier chapitre, section 1.1.4.

### 1.2.5 Lillian F. Schwartz aux Bell Labs. Un mariage de l'image animée par ordinateur et de la musique informatique



Fig. 26 : Lillian F. Schwartz utilisant un joystick de contrôle, en 1974  
© Center for Experimental Humanities

Lillian F. Schwartz est une plasticienne qui, à la fin des années 1960, va s'intéresser aux technologies informatiques. Elle a produit un nombre important de films et de pièces, mais son travail novateur et singulier reste encore relativement méconnu. L'analyse de son travail contribue également à éclairer le fonctionnement d'un des lieux d'émergence des arts par ordinateur, les Bell Labs<sup>137</sup>. Nous nous concentrerons ici sur trois de ses films dont les structures démontrent la singularité de ses approches liant l'image animée et musique.

Lillian F. Schwartz rencontre l'ingénieur Ken Knowlton, des Bell Labs, lors d'une présentation d'une pièce interactive à l'EAT<sup>138</sup>. Elle obtient une bourse AT&T en 1969. Knowlton commence à collaborer avec elle (et avec d'autres artistes également). Pendant plus de 20 ans, elle mènera un travail d'exploration artistique sur les relations entre ordinateur et image animée, sans jamais avoir de statut officiel<sup>139</sup>.

<sup>137</sup> Laboratoires de la société Bell Telephone, créés en 1879.

<sup>138</sup> Experiments in Art and Technology, voir section plus haut.

<sup>139</sup> Max Mathews s'était arrangé administrativement pour qu'elle puisse venir et circuler librement.

Peu après son arrivée, elle devient proche de Max Mathews, pionnier de la musique informatique aux Bell Labs, qui l'encourage à poursuivre ses expérimentations en lui faisant rencontrer les compositeurs en résidence. Selon Olivier Baudouin, près d'un tiers de la production musicale générée par ordinateur entre 1970 et 1977<sup>140</sup> sera utilisé pour ses films.

### ***Le rôle de Kenneth Knowlton et la figure de l'artiste-programmeur***

Kenneth Knowlton est une figure centrale dans le développement des premiers films programmés aux Bell Labs. Ingénieur de formation, il va commencer à créer des ponts entre l'art et la programmation. Il est, avant l'heure, un programmeur créatif au sens où l'entendra plus tard le designer interactif John Maeda et ses « légataires » avec le terme *Creative Coding* (Code de création). Il développe pour les Bell Labs l'un des tout premiers langages d'animation 2D, le BEFLIX<sup>141</sup>.

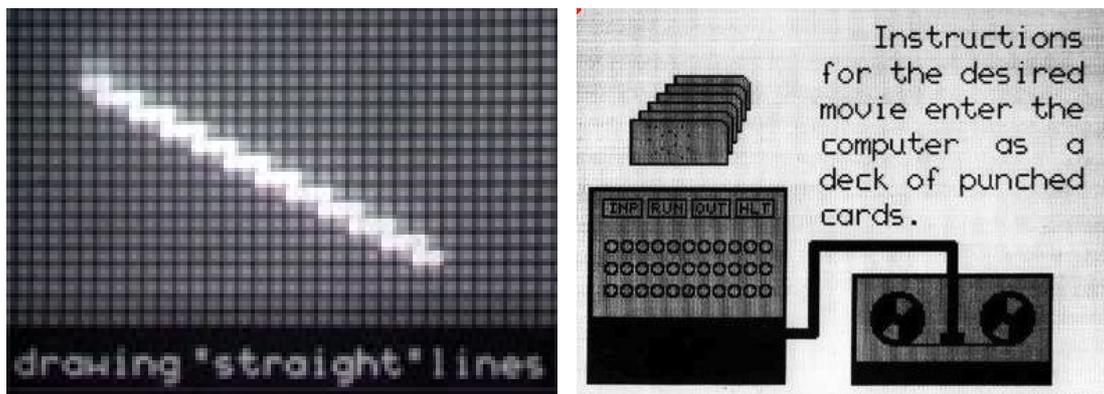


Fig. 27 : Captures d'un film 16 mm de documentation du BEFLIX, réalisé par Ken Knowlton en 1963. Ken Knowlton, *A Computer Technique for the Production of Animated Movies* (1963) © HistoryofInformation.com

140 Olivier Baudouin, *Pionniers de la musique numérique*, Sampzon, Delatour, coll. « Pensée musicale », 2012, p. 135.

141 Contraction de Bell flicks.

L'esthétique graphique induite par celui-ci évoque, pour le cinéaste et théoricien Malcolm Le Grice<sup>142</sup>, l'écran d'épingles d'Alexandre Alexeïeff et Claire Parke<sup>143</sup>. BEFLIX repose en effet sur une matrice, de 252 x 184 pixels.

Puis des artistes vont être invités à expérimenter sur ce nouvel outil qu'est l'ordinateur. Les témoignages de Knowlton à cet égard sont passionnants car il pointe différentes typologies de profil d'artiste. Knowlton pensait que ces artistes allaient s'initier à la programmation, mais la situation se révèle plus complexe. Au-delà des ego de chacun, certains ne veulent pas s'essayer à la programmation et sollicitent l'assistant technique de Knowlton. Pour Stan Vanderbeek, il développe une extension (ce que l'on appellerait aujourd'hui un *plugin*) qui crée une couche de programmation plus accessible (les instructions informatiques sont plus explicites, plus orientées et moins nombreuses), ceci permettant à l'artiste de programmer à minima. D'autres vont se former à la programmation et devenir quasiment autonomes, consultant uniquement Knowlton pour des cas complexes ; ils vont contribuer à faire émerger la figure de l'artiste-programmeur. Ces approches différentes, celles du programmeur et de l'artiste, sont toujours d'actualité. L'artiste en contact avec le numérique programme parfois lui-même, peut aussi commander un programme dédié à un développeur, ou bien travailler directement avec des logiciels clé en main. Il semble — et c'est l'expérience de Knowlton — que les deux collaborateurs gagnent à essayer de comprendre comment chacun travaille, afin de créer une dynamique mutuelle. Plus l'artiste comprend les rouages du numérique, plus il est à même d'en détourner les principes et d'avoir un regard critique. Le code, et par extension ses algorithmes sous-jacents, en est l'un des mécanismes intrinsèques.

### **Le langage EXPLOR**

Ken Knowlton crée spécialement pour Lilian Schwartz un nouveau langage, EXPLOR<sup>144</sup>, proche de BEFLIX, qui va simplifier sa programmation. Cela va lui permettre de pro-

---

142 Yann Beauvais (dir.), *Malcolm Le Grice. Le temps des images*, Dijon, Les presses du réel, 2015, p. 87.

143 *Une Nuit sur le Mont Chauve* (1933) est leur premier film à utiliser cette technique. Il s'agissait d'une « [...] trame de pointes traversant une surface blanche, qui, éclairées obliquement, donnent naissance à autant d'ombres portées plus ou moins longues selon la saillie des épingles sur la surface blanche » (CNC, « L'écran d'épingles Alexeïeff/Parker », juin 2021, [https://www.cnc.fr/cinema/actualites/lecran-depingles-alexeieff%E2%80%94parker\\_1219853](https://www.cnc.fr/cinema/actualites/lecran-depingles-alexeieff%E2%80%94parker_1219853), consulté le 30/01/23).

144 Contraction de EXplicitily Provided 2D Patterns, Local Neighbourhood Operations And Randomness.

duire des compositions animées expérimentales mêlant prises de vues réelles traitées par ordinateur et générations de motifs abstraits qui adoptent l'esthétique induite<sup>145</sup> par BEFLIX. Pour ses bandes-son, elle travaille avec les compositeurs invités aux Bell Labs, en investissant leur musique générée par ordinateur.

Après plusieurs expérimentations sur l'image fixe, Schwartz se consacre à ce qui deviendra son premier film, *Pixillation* (1970)<sup>146</sup>, d'une durée de 4 minutes. La première composante du film est sa partie réalisée par ordinateur, assemblage animé de gros pixels évoluant dans le temps, réalisée avec EXPLOR. Les premières versions d'EXPLOR étant très chronophages, elle ne réussit à coder et produire que quelques secondes en deux mois<sup>147</sup>. Elle décide donc de compléter son film avec des séquences filmées image par image : des compositions plastiques qu'elle produit elle-même<sup>148</sup>, beaucoup plus rapidement. Elle venait également d'acquérir pour son logement personnel un Moviola<sup>149</sup>, ce qui lui permettait de vérifier ce que donnait le film avec le son.

Les parties peintes sont filmées avec une caméra 16 mm BOLEX. Elle entremêle au final trois natures d'images : image noir et blanc générée par ordinateur, images peintes et des minéraux en croissance accélérée. Pour elle, les images informatiques seules étaient trop pauvres. En peignant des formes carrées similaires à celles de l'ordinateur, et en jouant sur les couleurs, elle pouvait produire plus de variété<sup>150</sup>.

---

145 De par ses limitations, le langage d'animation, pour un regard contemporain, produit en tout cas certains effets pixelisés reconnaissables.

146 Une version est disponible en ligne sur le Vimeo de Lillian F. Schwartz, <http://lillian.com/1970-pixillation-4-min/>, consulté le 30/01/23.

147 Rebekah Rutkoff, « Painting by numbers: the art of Lillian Schwartz », Art Forum, octobre 2016, <https://www.artforum.com/print/201608/painting-by-numbers-the-art-of-lillian-schwartz-63453>, consulté le 30/01/23.

148 D'abord aux Bell Labs, puis chez elle.

149 Machine de visionnage au format réduit, sans projection, de films sur pellicule. Inventé dans les années 1920, elle permet également d'entendre le son synchronisé.

150 Chris Garcia, interview de Lillian F. Schwartz, « Oral History of Lillian F. Schwartz », août 2013, <http://archive.computerhistory.org/resources/access/text/2016/06/102746737-05-01-acc.pdf>, consulté le 30/01/23.

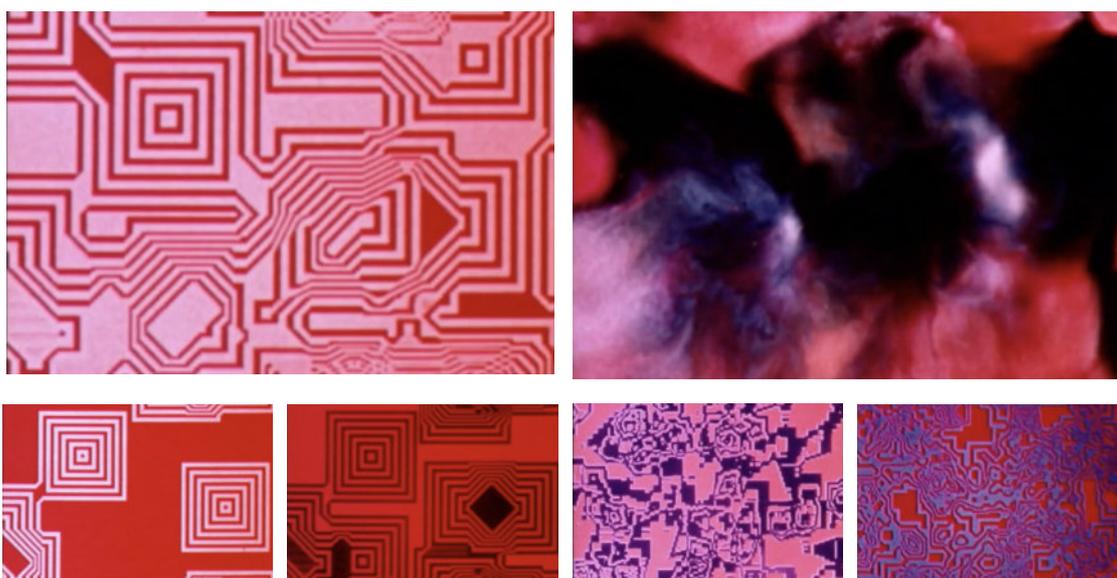


Fig. 28 : Lillian Schwartz, *Pixillation* (1970)  
avec l'assistance de Kenneth Knowlton  
© Site web de Lillian Schwartz, lillian.com

Le film est en général présenté<sup>151</sup> avec une bande-son électronique (au Moog<sup>152</sup>) du compositeur-musicien Gershon Kingsley<sup>153</sup>. Il existe cependant quatre versions différentes de ce même film, chacune utilisant la même composante visuelle, mais une bande-son différente. L'idée est de tester plusieurs « ambiances<sup>154</sup> ». L'une d'entre elles, celle de Richard Moore, est composée à l'aide du système GROOVE<sup>155</sup> qu'il a inventé en 1969 avec Max Mathews. Il s'agit d'un ordinateur interfacé avec un synthétiseur analogique et couplé à une enceinte. L'ordinateur convertissait et stockait tout ce qui était composé, et tout ce qui était stocké pouvait donc être rappelé, manipulé, et sauvegardé<sup>156</sup>.

Les autres bandes-son sont celles de Frank Lewin (compositeur et enseignant en musique de film à Yale), pour clavecin, ondioline<sup>157</sup> et célesta, et un autre accompagnement alternatif de Kingsley au Moog. La première bande-son, celle de Moore, semble

151 Voir sur le site de l'artiste, <http://lillian.com/films/>, consulté le 30/01/23.

152 Un synthétiseur modulaire.

153 Compositeur du morceau *Popcorn* (1969).

154 Lillian F. Schwartz, *The Computer Artist's Handbook: Concepts, Techniques, and Applications*, New York, WW Norton & Company, 1992, p. 194.

155 Contraction de Generated-Realtime-Operations-On-Voltage-Electronics.

156 Lillian F. Schwartz, *op. cit.*, p. 185.

157 Inventé en 1941 par Georges Jenny. Instrument électronique à clavier. Ancêtre du synthétiseur, assez proche dans son fonctionnement des ondes Martenot.

avoir été refusée en partie parce que la copie était pauvrement transférée sur film, mais ces versions circulaient dans les milieux artistiques de l'époque<sup>158</sup>.

Lillian Schwartz inscrit explicitement son travail dans la lignée des films des années 1920-1930 de Viking Eggeling et d'Oskar Fischinger, en pointant le développement d'un langage<sup>159</sup>. En 1972, deux films importants voient le jour : *Mutations* et *Goolplex*. *Mutations*<sup>160</sup> est réalisé à partir de la bande-son préexistante<sup>161</sup>, éponyme, du compositeur Jean-Claude Risset<sup>162</sup>. Cette pièce est connue pour l'utilisation de la gamme de Shepard, illusion auditive qui donne la sensation d'entendre une chute/montée sans fin, à la manière des escaliers d'Escher. C'est aussi l'un des premiers morceaux à utiliser la synthèse FM.

Par ailleurs, Lillian Schwartz s'intéresse aux travaux des scientifiques. Déjà en 1971, elle utilisait pour *UFO's* des chutes de films scientifiques. Pour *Mutations*, elle filme les expérimentations de Don White, qui s'intéresse à l'époque à des lasers holographiques à 360°. Les images obtenues sont proches des *Lumia* de Thomas Wilfred, ou bien des images produites par John Whitney dans *Catalog* avec ses ordinateurs analogiques. Après l'abstraction géométrique, ces images fantomatiques ont une sorte de congruence naturelle avec la bande-son, formant une sorte de contrepoint naturel,

---

158 Olivier Baudoin, *op. cit.*, p. 136.

159 Lillian F. Schwartz, *op. cit.*, p. 152.

160 Une version est disponible en ligne sur le Vimeo de Lillian F. Schwartz, <http://lillian.com/films/>, consulté le 30/01/23.

161 Olivier Baudoin, *op. cit.*, p. 103-115.

162 Réalisée avec le logiciel MUSIC V.

comme le décrivait Schaeffer<sup>163</sup>, à propos de l'une de ses expérimentations audiovisuelles au service de la recherche de l'ORTF<sup>164</sup>.

Lillian Schwartz poursuivra ses recherches avec Risset, qui composera pour elle la musique de deux autres films, *Olympiad* et *Papillons*. *Mutations* est présenté au festival de Cannes en 1974.

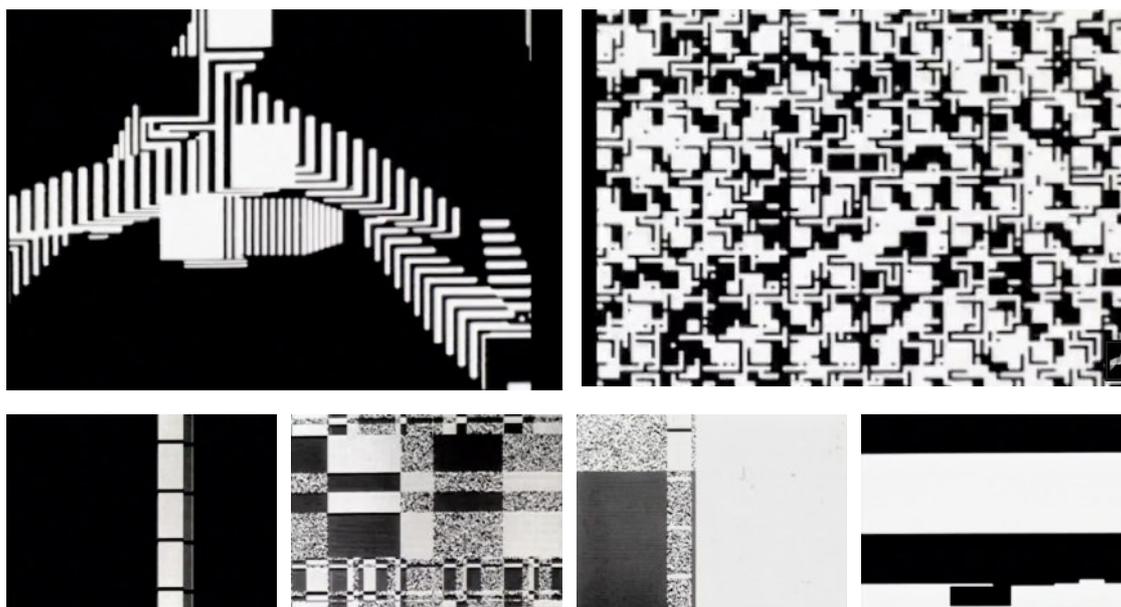


Fig. 29 : Lillian F. Schwartz, *Googolplex* (1972)  
avec l'assistance de Kenneth Knowlton  
© Site web de Lillian Schwartz, lillian.com

163 Il s'agit de l'œuvre audiovisuelle *Étude aux allures* (1960), réalisée en collaboration avec le plasticien Raymond Hains. Celui-ci avait produit des séquences filmées réalisées avec des lentilles cannelées, produisant une sorte d'abstraction graphico-liquide en évolution. Schaeffer et les autres musiciens du laboratoire reconnaissent dans ces images un équivalent de l'« allure », l'une des caractéristiques liées à l'entretien d'un son décrites dans le *Traité des objets musicaux*. Schaeffer propose à Hains d'utiliser cette pièce de musique concrète *Étude aux allures* (1958) — qui donnera également son nom au film — comme musique. Schaeffer en décrit ainsi la relation audio-visuelle : « [...] ce qui fut le plus remarquable, c'est l'aisance, précisément, d'un contrepoint qui pouvait se dérouler fort librement, dès que la structure des images et des sons étaient en parenté. Aucun besoin d'ajuster numériquement battements ou mouvements. C'est l'esprit qui comptait, et non la lettre, moyennant un certain respect du "scénario", du dialogue sans parole qui s'établissait de lui-même entre les événements Son et Image » (Pierre Schaeffer, « Le contrepoint du son et de l'image », in *Cahiers du cinéma*, n° 108, juin 1960, p. 21).

164 Pierre Schaeffer est connu pour ses travaux sur les objets sonores, mais il a également, dans le cadre du Service de la recherche de l'ORTF, contribué entre 1960 et 1974 à faire se rencontrer des plasticiens cinéastes (Piotr Kamler ou Peter Foldes par exemple) et des compositeurs de musique concrète (tels que Luc Ferrari ou Bernard Parmegiani), dans l'idée d'étendre sa recherche aux objets audiovisuels. De nombreuses créations audiovisuelles expérimentales sont nées de ces croisements, à la manière des travaux de Lillian F. Schwartz aux Bell Labs. Philippe Langlois a très bien retracé cette période méconnue de l'ORTF dans plusieurs textes, notamment au sein de son ouvrage *Les cloches d'Atlantis* (Philippe Langlois, *Les cloches d'Atlantis : musique électroacoustique et cinéma, archéologie et histoire d'un art sonore*, Paris, Éditions MF, coll. « Répercussions », 2012).

Pour *Googolplex* (1972)<sup>165</sup>, film entièrement en noir et blanc, la réalisation d'images est devenue un peu moins complexe, la totalité des images du film est générée par ordinateur. La bande sonore, en revanche, puise sa source dans les musiques percussives africaines. Elle cale précisément les images sur le rythme de la musique<sup>166</sup>, jouant sur les effets de *flicker* (clignotement). Cet effet de scintillement persiste tout au long du film, suivant son rythme.

Le langage visuel est primaire, des abstractions graphiques en animations syncopées. Il n'y a plus ici d'images abstraites provenant du réel. Mais le film relève d'une esthétique très contemporaine pour l'époque. Il mêle des formes de type BEFLIX/EXPLOR, avec des superpositions de trames de points (proches de la neige visuelle sur une télévision lorsque le signal n'est pas capté), avec des aménagements d'espaces blancs ou noirs<sup>167</sup>. Il existe trois versions de ce film, avec un agencement sonore différent. La séquence débute par ce qui semble être un point d'interrogation, puis évolue en formes abstraites.

En 1974, elle se consacre à *Galaxies*, basé sur des images de Frank Hohl à la NASA. Elle commence à travailler la bande-son avec Richard Moore, puis propose également à Vladimir Ussachevsky (de passage pour voir Max Mathews) de travailler sur une version alternative. Elle a finalement combiné la composition digitale de Moore et celle, analogique, d'Ussachevsky pour le film<sup>168</sup>.

Le travail de Lillian Schwartz explore donc des formes abstraites par ordinateur mais aussi du *Found Footage* d'imagerie scientifique, des films de laser, des peintures, et fait cohabiter, dialoguer des musiques électroniques et informatiques « abstraites ».

En mariant ses images avec de la musique informatique, Schwartz explore l'adéquation entre certaines caractéristiques de l'image et du son, qui ne sont ni de l'ordre de la synchronicité ni de l'ordre du signifié. On peut ici se rapprocher du concept de substance, proposé par le compositeur Michel Fano, qui permettrait d'aider à décrire des

---

165 Une version est disponible en ligne sur le Vimeo de Lillian F. Schwartz, <http://lillian.com/1972-googolplex-5-12-min/>, consulté le 30/01/23.

166 Lillian F. Schwartz, *op. cit.*, p. 203.

167 Ceux-ci ne sont pas sans rappeler *Rythme 21* (1921) d'Hans Richter et *Noisefields* (1974) des artistes vidéo Steina et Woody Vasulka, bien qu'ayant un rythme plus effréné, tribal.

168 Lillian F. Schwartz, *op. cit.*, p. 195.

situations dans lesquelles des images non représentatives et des sons non iconiques seraient en relation, notamment dans des formes audiovisuelles non traditionnelles, qu'il nomme « Filmusique ».

[...] Cet ensemble de fonctions se fonde sur un rapport, du son à l'image, déterminé par les axes du temps, de la substance, et de la signification. L'application d'un son sur une image révélera aussitôt la relation au temps : synchronisme ou non-synchronisme ; la relation à la substance : congruence ou non-congruence ; la relation à la signification : référent ou non-référent<sup>169</sup>.

Si Schwartz est pionnière dans l'art de l'ordinateur, elle est également une actrice importante des arts audiovisuels, en plaçant systématiquement l'expérimentation avec des musiques nouvelles au centre de ses préoccupations, en donnant une place aussi importante à l'image et au son.

### 1.2.6 Une esthétique programmée

Comme nous l'avons vu, cette première période de l'art par ordinateur est celle de l'émergence de la notion de « programme ». Celle-ci est toutefois antérieure à l'utilisation concrète de l'ordinateur par les artistes. En effet, alors que l'art visuel par ordinateur va émerger au début des années 1960, l'utilisation du terme — du latin *programma*, qui signifie « ce qui est écrit à l'avance », et qui revêt avant les années 1960 le sens d'un programme élaboré en vue d'une diffusion, par exemple radiophonique ou télévisuelle<sup>170</sup> — s'insinue dans le champ des arts plastiques et visuels par le biais de différents intermédiaires. Parmi eux le peintre et sculpteur François Morellet évoque une expérience de peinture « menée à partir d'éléments contrôlables en progressant systématiquement suivant un programme<sup>171</sup> ». L'écrivain et essayiste Umberto Eco, dans un texte du catalogue de l'exposition *Arte Programmata* en 1962, décrit ainsi la nature des

169 Michel Fano, « Musique et film : filmusique », In *Vibrations*, n° 4, « Les musiques des films », 1987, p. 27.

170 Francine Dagenais, « Programmation : l'écriture, le système, le langage, le son, le rythme », in *Revue Inter-médialités/Intermediality*, n° 13, printemps 2009, p. 17–37, <https://www.erudit.org/en/journals/im/2009-n13-im3872/044038ar/>, consulté le 30/01/23.

L'autrice précise que l'omission des voyelles dans le titre fait allusion à *HPSCHD* (1967 -1969), une œuvre pionnière de John Cage et de Lejaren Hiller, qui utilise l'ordinateur.

171 François Morellet, « Pour une peinture expérimentale programmée », article publié dans la plaquette du Groupe de recherche d'art visuel, Paris, Galerie Denise René et le GRAV, 1962. Accessible en ligne : <https://francois-morellet.fr/2021/06/01/pour-une-peinture-experimentale-programmee-1962/>, consulté le 30/01/23.

travaux présentés : « Il ne sera donc pas impossible de délimiter avec la pureté linéaire d'un programme mathématique, des champs des cas possibles dans lesquels peuvent se vérifier des processus casuels. C'est ainsi que nous aurons une singulière dialectique entre cas et programmes, entre mathématique et hasard, entre conception planifiée et libre acceptation de ce qui arrive... Nous pouvons parler d'art programmé<sup>172</sup> ». Dans *l'Œuvre ouverte*, Eco détaillait également une telle nature d'œuvre contemporaine, « en mouvement » :

En somme, l'auteur offre à l'interprète une œuvre à achever. Il ignore de quelle manière précise elle se réalisera, mais il sait qu'elle restera son œuvre ; au terme du dialogue interprétatif se concrétisera une forme organisée par un autre, mais une forme dont il reste l'auteur. Son rôle consiste à proposer des possibilités déjà rationnelles, orientées et dotées de certaines exigences organiques qui déterminent leur développement<sup>173</sup>.

Il s'appuie, pour sa démonstration, sur des exemples issus du champ de la musique contemporaine, à travers les pièces de quatre compositeurs (Stockhausen, Berio, Pousseur et Boulez) dont les exécutions offrent aux musiciens des degrés de liberté dans l'interprétation. Citons également le designer Karl Gerstner, qui dans son ouvrage *Designing Programmes*<sup>174</sup> (1964), évoque une méthode de design graphique fondée sur des processus combinatoires. À travers de tels exemples, on comprend que la musique ou les arts visuels peuvent indifféremment être concernés par ces logiques.

La première période de l'art par ordinateur a montré que ces méthodes devenaient partie intégrante du processus de l'artiste. En 1971, le travail de Manfred Mohr<sup>175</sup> était au centre de la première exposition monographique d'art par ordinateur en France, au MAM. Elle était intitulée *Computer Graphics — Une Esthétique Programmée*. Dans le catalogue<sup>176</sup> on pouvait lire ce texte :

172 Umberto Eco, *Arte programmata*, catalogue de l'exposition « Arte programmata », 1962, cité et traduit par Frédéric Migayrou, Frédéric Migayrou, « Conjectures des langages formels », in *Coder le monde*, *op. cit.*, p. 70.

173 Umberto Eco, *L'Œuvre ouverte*, Paris, Seuil, coll. « Points », 1965, p. 34.

174 Karl Gerstner, *Designing Programmes. Programme as Typeface, Typography, Picture, Method*, Salestein, Niggli Verlag, 1964.

175 Il participera à la création du groupe « Art et informatique » à l'Université de Vincennes en 1968.

176 Celui-ci est accessible en ligne sur le site de Manfred Mohr : <http://www.emohr.com/paris-1971/catalog/layoutcatalog71.html>, consulté le 30/01/23.

Or c'est bien ce que produit un programme de Mohr à chaque exception : un membre d'une famille de graphismes identiques et différents. Nous nous trouvons face à un paradoxe qui peut avoir des conséquences théoriques intéressantes. Il nous permet de définir de façon rigoureuse l'égalité esthétique entre différents tableaux obtenus à partir du même programme. L'activité créatrice est intervenue une fois pour toutes en introduisant les éléments graphiques et les règles de composition. Une séparation claire est établie entre ce que l'artiste considère comme essentiel (les règles du jeu) et contingent (les décisions aléatoires)<sup>177</sup>.

Cette esthétique programmée est caractérisée par ce que le cinéaste expérimental et théoricien Malcolm Le Grice<sup>178</sup> appelle « programmabilité ». Celle-ci représente une troisième étape technologique dans l'histoire des arts, après ce qu'il nomme les « médias directs de la main et du corps » (peinture, musique, sculpture, danse), puis les « médias de reproduction mécanique » (presse imprimée, photographie, film, phonographe) :

[l'ordinateur] est un système intelligent nécessitant des instructions sur le travail qu'il doit accomplir et la manière dont il doit dialoguer avec d'autres machines. Ces instructions constituent des programmes [...]. Une caractéristique majeure du programme est que ses productions changent selon les données ou les procédures choisies. Le programme est un ensemble d'instructions permettant de manipuler des données d'une façon particulière et pouvant inclure la réponse à un nouvel apport de données ou une nouvelle information quant à la manière dont elles peuvent être manipulées. Par conséquent, les œuvres produites se conforment à un modèle cohérent formé par le programme, mais chaque version particulière de l'œuvre peut présenter des différences. Ces différences font partie intégrante du travail et peuvent être basées sur une gamme de facteurs ou de stratégies inhérentes au programme<sup>179</sup>.

177 Andrés Zucker, « Art et ordinateur. Trois réflexions sur une esthétique programmée : section II », in Manfred Mohr, *Computer Graphics - Une Esthétique Programmée*, catalogue d'exposition, Paris, ARC/Musée d'Art Moderne de la Ville de Paris, 1971, p. 30.

178 Cinéaste expérimental important, il a aussi été l'un des premiers artistes britanniques à travailler avec l'ordinateur avec le film muet *Your Lips* (1970), puis avec plusieurs films postérieurs mettant en jeu des relations numériques entre l'image et le son, dont nous parlerons dans cette section. Dans ses ouvrages *Abstract film and beyond* (1977) il analyse les formes du cinéma expérimental abstrait, et dans *Experimental Cinema in the Digital Age* (2002), l'émergence du numérique dans le cinéma expérimental.

179 Malcolm Le Grice, « Cinéma numérique et film expérimental. Continuités et discontinuités », 1999, in Yann Beauvais (dir.), *Malcolm Le Grice. Le temps des images*, op. cit., p. 216.

La notion de programmabilité est aussi fondamentale chez Kittler, qui prône d'ailleurs l'apprentissage pour tous d'au moins deux langages de programmation<sup>180</sup>, en complément de l'alphabet et des nombres mathématiques. Elle existe également chez Hillaire et Couchot qui utilisent quant à eux le terme programmabilité, c'est-à-dire le fait que les technologies numériques « se réduisent à des programmes informatiques capables d'être traités automatiquement par la machine ordinateur<sup>181</sup> ».

Dans le processus théorisé par Le Grice, au-delà de l'éventail des différentes versions ou résultats qui viennent faire œuvre, la programmabilité induit une capacité de réponse en temps réel du programme, une réactivité. Si cette réactivité se produit à la suite de l'intervention d'un utilisateur, on parlera d'interactivité ; si celle-ci se produit de manière autonome et endogène, dans des temps définis par le programme, on parlera de générativité. Au début des années 1980, avec le développement du traitement en temps réel, se développent de nouvelles formes jouables. Elles oscillent, comme le formule Douglas E. Stanley, entre la « variabilité manuelle de l'utilisateur et l'autonomie de la simulation construite dans la machine<sup>182</sup> » et produisent « un jeu entre interactivité et générativité<sup>183</sup> ». Le programme devient instrument : « Il faudrait entendre dans le mot *instrument* les multiples sens qu'il peut prendre : instrument musical, instrument technique ou de contrôle, instrument à fin particulière. L'informatique met en jeu ses trois définitions, à cause de sa structure algorithmique, programmée<sup>184</sup> ».

---

180 Frédérique Vargoz, « Friedrich Kittler. Y a-t-il quelqu'un dans la machine ? », in *MCD #66*, « Machines d'écritures », mars/mai 2012, <https://www.digitalmcd.com/friedrich-kittler/>, consulté le 30/01/23.

181 Edmond Couchot et Norbert Hillaire, *op. cit.*, p. 26.

182 Douglas E. Stanley, « This is Fun! », in Jean-Louis Boissier et Daniel Pinkas, *Jouable. Art, jeu et interactivité*, Genève, Centre pour l'image contemporaine/ Haute école d'arts appliqués HES/ École nationale supérieure des arts décoratifs/ Ciren, Université Paris 8, 2004, p. 112.

183 *Ibid.*, p. 112.

184 *Ibid.*, p. 114.

## 1.3 *Inventer ses outils : des interfaces artistiques pour un art audiovisuel*

### 1.3.1 *Live Visuals. Générer le visuel en temps réel*

Très tôt déjà, dans les années 1920, avec l'invention d'un instrument tel que le Theremin<sup>185</sup>, il a été possible de générer des sons nouveaux, par la médiation d'une machine et du geste de l'interprète, en temps réel. Le développement de la musique électronique, puis des synthétiseurs modulaires tels que le Moog ou le Buchla — plus expérimental — ont ensuite permis, entre autres formes musicales, le développement de performances de musique électronique Live (*Live Electronic Music*<sup>186</sup>).

L'image électronique générative et manipulable arrive plus tard<sup>187</sup>. Certains principes du *Live* préexistent toutefois déjà dans les années 1960. Pour accompagner certains concerts de musique pop, des collectifs d'art visuel se sont formés. Les *Light Shows*, projetés derrière le groupe de musique, étaient produits en direct. Les membres du collectif mélangeaient des liquides colorés dans des récipients et les filmaient en gros plan. Le tout était retransmis en direct sur l'écran de la scène. C'est toutefois principalement avec l'art vidéo qu'une image abstraite peut être générée et manipulée en temps réel. L'invention de synthétiseurs vidéo<sup>188</sup>, instruments visuels à claviers similaires aux synthétiseurs audio, a permis à l'image de devenir aussi malléable que la matière sonore l'était devenue. Le fait que le synthétiseur vidéo s'inspire des procédés musicaux n'est pas anodin : une grande majorité des pionniers de l'art vidéo étaient au

185 Ou thérémine, inventé par Leon Theremin vers 1920. Il s'agit d'une machine à deux antennes. L'interprète fait varier l'amplitude et la tonalité du son en s'approchant ou s'éloignant des antennes, sans aucun contact. Il existait même une expérimentation dans laquelle le Theremin était relié électriquement à un tube de néon, qui projetait des flashes lumineux en fonction des fréquences (Laurent de Wilde, *Les fous du son. D'Edison à nos jours*, Paris, Grasset, p. 126).

186 Le musicien Gordon Mumma, à la fin des années 1950, en est l'un des pionniers. Il met lui-même au point ses propres dispositifs électroniques, qu'il nomme « Cybersonic ».

187 Même si les projections de lumière avec les orgues colorés — tels que le Clavilux de Thomas Wilfred — ont permis des manipulations en direct sous forme de projection en grand format, il ne s'agissait pas encore d'une image-média écranique.

188 Parmi ces premières expériences, celle du Paik-Abe Video Synthesizer. Entre 1969 et 1970, Nam June Paik et le technicien Shuya Abe, avec le soutien d'une station de télévision, conçoivent un synthétiseur vidéo. Celui-ci leur permet de mixer et de manipuler en direct des images. Cet outil était destiné au live et devait pouvoir être utilisé aussi simplement qu'un piano. Le Paik-Abe Video Synthesizer est utilisé par Paik en 1970, durant l'émission de télévision « Videocommune » sur WGBH, avec l'équipe de la station TV. Ils improvisent pendant 4 h, avec la musique des Beatles comme support musical. L'improvisation avec des images est ainsi rendue possible (Dieter Daniels, « Sound & Vision in Avant-garde & Mainstream », 2004, [http://www.medienkunstnetz.de/themes/image-sound\\_relations/sound\\_vision/13/](http://www.medienkunstnetz.de/themes/image-sound_relations/sound_vision/13/), consulté le 30/01/23).

départ musiciens ou gravitait dans le monde musical. Citons Bill Viola, Nam June Paik, ou encore le duo Steina et Woody Vasulka<sup>189</sup>. Ces derniers rappellent le rôle fondamentale de la musique à ce moment clé :

Tout comme la musique instrumentale a été le modèle opératoire du cinéma abstrait, et elle le reste encore dans de nombreuses expériences d'aujourd'hui, le son électronique a été le modèle opératoire de la vidéo<sup>190</sup>.

Nous ne pouvons pas décrire toutes les machines et principes qui émergent au sein de cette effervescence électronique et vidéo<sup>191</sup>, mais nous pouvons pointer une performance historique de l'époque : *Illuminated Music* (1972-1973)<sup>192</sup> de Stephen Beck, artiste vidéo, et Warner Jepson, musicien.

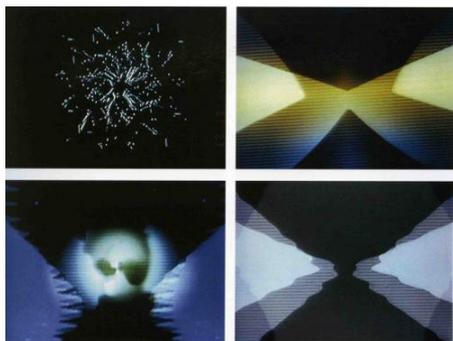


Fig. 30 : Stephen Beck et Warner Jepson, *Illuminated Music* (1973)  
© Ubu.com



Fig. 31 : Stephen Beck « jouant » de son Direct Video Synthesizer  
© mmmmm.ch / Electronic Arts Intermix

Il semble s'agir en effet de la toute première performance audio-vidéo, dans laquelle à la fois le son (généralisé par le mythique synthétiseur modulaire Buchla<sup>193</sup>) et l'image

189 Woody et Steina Vasulka sont des pionniers de l'art vidéo. Nous reparlerons de leur processus de transformation des images en sons et des sons en images dans le prochain chapitre.

190 Marcella Lista, « Empreintes sonores et métaphores tactiles », in Sophie Duplex, Marcella Lista (dir.), *op. cit.*, p. 74.

191 En complément de nos références déjà citées, l'ouvrage suivant, Steve Gibson, Stefan Arisona, Donna Leishman, Atau Tanaka, *Live Visuals : History, Theory, Practice*, Abingdon-on-Thames, Taylor & Francis Ltd, 2022, que nous n'avons pas encore pu consulter (car sorti tout récemment), semble couvrir cette période, de même qu'il paraît traiter de problématiques tout à fait proches des nôtres dans cette première partie.

192 Une version est disponible en ligne sur Ubu.com, [https://ubu.com/film/beck\\_illuminated.html](https://ubu.com/film/beck_illuminated.html), consulté le 30/01/23.

193 Utilisé par le compositeur Morton Subotnick, en 1963, dans sa pièce *Silver Apples on the Moon*.

(produite à l'aide du *Direct Video Synthesizer*, inventé par Stephen Beck lui-même, et qui permet de générer des images sans l'aide d'une caméra) sont générés et dialoguent en temps réel, électroniquement<sup>194</sup>. Comme le souligne Dieter Daniels<sup>195</sup>, d'autres synthétiseurs vidéo, souvent basés sur des synthétiseurs audio, font leur apparition à cette période. Si ceux-ci permettent une forme d'improvisation, ils sont encore très rares sur scène. En 1973, les deux artistes font une mini-tournée et présentent des performances réunies sous le nom d'*Illuminated Music*. Même si tout est encore ici analogique, il y a un rapport d'équilibre entre l'image et le son, considérés tous deux comme un matériau équivalent. Dans le dispositif de Beck et de Jepson, les circuits de l'audio et du visuel ne sont pas connectés. Il s'agit donc bien d'improvisation et d'accordages mutuels. L'image étant abstraite, et la musique très abstraite également, il y a véritablement là le développement d'un langage audiovisuel.

Dans le champ numérique, la première connexion en temps réel de formes graphiques générées et de sons générés semble avoir été réalisée par la compositrice Laurie Spiegel.



Fig. 32 : Laurie Spiegel  
© WFMU

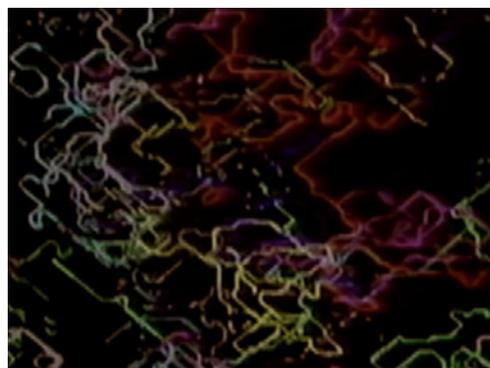


Fig. 33 : Laurie Spiegel, *Voyages* (1977)  
Réalisé avec VAMPIRE  
© Laurie Spiegel YouTube

194 Nous l'avions décrite lors d'une conférence en duo avec Gilles Alvarez, directeur du festival (devenu biennale) NémO, dédiée aux arts numériques contemporains (Gilles Alvarez et Benoît Montigné, « Musique live/Images live, histoires de rencontres », dans le cadre de la Biennale NémO, Cité de la Musique, 10 octobre 2015).

195 Sandra Naumann, « The Expanded Image: On the Musicalization of the Visual Arts in the Twentieth Century », in Dieters Daniels, Sandra Naumann (dir.), *See This Sound, Audiovisuology 2. Essays. Histories and Theories of Audiovisual Media and Art*, Köln, Walther König, 2011, p. 86.

Vers 1974, aux Bell Labs, Laurie Spiegel apprend à utiliser le système GROOVE<sup>196</sup> de Max Mathews, qui permet de manipuler numériquement du son analogique généré en temps réel. Le vocabulaire sonore qu'elle utilise, par exemple dans *Appalachian Grove* (1974) est volontairement très minimal. Autour d'un même ensemble de notes<sup>197</sup>, elle y explore des variations de spatialisation et de durée. Inspirée par l'émergence de l'art graphique par ordinateur — les Bell Labs en sont l'un des foyers d'émergence —, elle a alors l'idée d'étendre son travail sonore en temps réel à l'image, en intégrant une composante graphique qui pourrait être produite en direct. Elle développe entre 1974 et 1976 le système *VAMPIRE* (Video And Music Program for Interactive Realtime Exploration/Experimentation), c'est-à-dire un système qui pourrait générer en temps réel à la fois la partie graphique et la partie sonore. Elle parvient à générer des formes graphiques en temps réel, mais elle n'arrive toutefois pas à stabiliser le dispositif en couplant l'image et le son directement<sup>198</sup>. Cela en partie parce que les ordinateurs dédiés à l'image et au son étaient dans des espaces séparés, et ne permettaient donc pas facilement une connexion directe. Cela était toutefois potentiellement réalisable.

Même si le système de Spiegel est resté inachevé, il s'agit là du début d'une nouvelle ère, celle de l'art génératif et interactif en temps réel. À ce titre, nous souhaitons dans les deux sections suivantes mettre en valeur deux œuvres. Dans la première, *Arbitrary Logic* (1984-86), le cinéaste et théoricien Malcolm Le Grice, utilise à la fois la générativité audiographique et l'interactivité, en temps réel, celle que Laurie Spiegel avait quasiment mise en œuvre. Dans la seconde, *Moon Drum* (1991), John Whitney utilise l'interactivité et la générativité dans un logiciel spécialement développé pour lui, pour composer des contrepoints audio-graphiques.

---

196 « Le système GROOVE, ou Generating Real-time Outputs Operations on Voltage-controlled Equipment (opérations en temps réel générées en sortie sur appareil à commandes en tension), possède en sortie quatorze convertisseurs numérique-analogique et offre au musicien autant de fonctions logicielles, contrôlées en temps réel par trois boutons rotatifs, quatre boutons à bascule, deux boutons-poussoirs et un clavier de type machine à écrire. Un levier de commande permet de se positionner dans le temps et de faire défiler l'affichage des valeurs en fonction du temps sur un écran cathodique » (Olivier Baudouin, *op. cit.*, p. 124).

197 Les quatre notes do, mi, sol et sib.

198 Un extrait de *Voyages* (1977), pièce audio-vidéo réalisée avec VAMPIRE est disponible en ligne sur le YouTube de Laurie Spiegel, mais sans le son lié, <https://www.youtube.com/watch?v=0Fy5iIFJENo>, consulté le 30/01/23.

### 1.3.2 Malcolm Le Grice, *Arbitrary Logic*. Un orgue à couleurs génératif

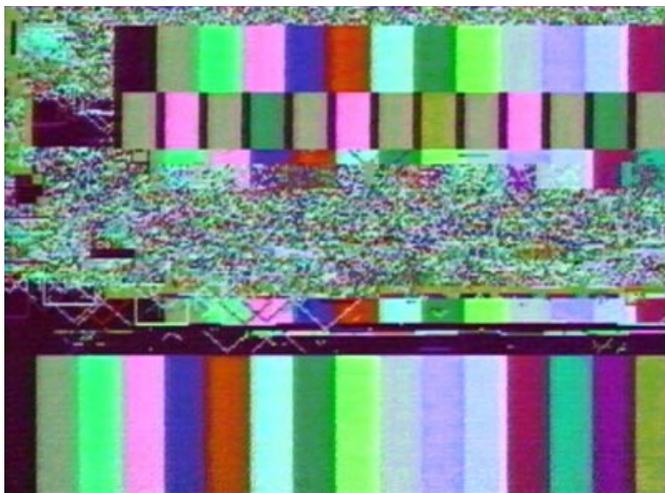


Fig. 34 : Malcolm Le Grice, *Arbitrary Logic* (1984-86)

Capture d'écran

© REWIND Artists' Video

Comme le souligne Philippe Langlois, dans beaucoup de ses films expérimentaux Malcolm Le Grice explore les relations entre image et son, à travers « l'exploration de l'interstice qui se glisse entre le son et l'image, entre l'image et l'image, entre le son et l'absence de son, dans cet intervalle où se situe une dimension poétique élémentaire d'un grand raffinement<sup>199</sup> ». Parmi eux *Berlin Horse* (1970), diptyque filmique dans lequel Le Grice travaille les refilmages successifs d'une séquence en boucle, celle d'un cheval tournant autour d'un poteau. La musique, de Brian Eno, est également une boucle, mais de musique électronique qui se déphase progressivement. Même si la musique est ajoutée postérieurement et n'est donc pas pensée selon une partition précise entre image et son, l'ensemble constitue un véritable objet audiovisuel qui met en jeu la figure dans deux modalités, à travers la confrontation de la boucle visuelle et de la boucle sonore. La même année, suite à une première expérience de performance utilisant un ordinateur<sup>200</sup>, et découvrant les films informatiques de John Whitney et de

199 Philippe Langlois, « La musique des images », in Yann Beauvais (dir.), *Malcolm Le Grice. Le temps des images*, op. cit., p. 288.

200 En 1969, il crée avec le programmeur et artiste sonore Alan Sutcliffe la performance Typo-Drama. Le texte et les instructions destinés aux acteurs avaient été produits par un programme informatique. La pièce est présentée à *Event One*, à Londres, en 1969, première exposition organisée par la Computer Arts Society ; Alan Sutcliffe était l'un des trois membres fondateurs (avec George Mallen and John Lansdown) de ce collectif, qui existe toujours, en 1968. C'est Alan Sutcliffe qui initie Malcolm le Grice à la programmation informatique. Des archives d'*Event One* sont accessibles à cette adresse : <https://computer-arts-society.com/event-one>, consulté le 30/01/23.

Stan Vanderbeek, il réalise au début des années 1970 *Your Lips*, film par ordinateur qui lui a demandé 9 mois de programmation en FORTRAN 4 pour seulement quelques secondes de film. Il ne réemploie le numérique que dans les années 1980, avec l'arrivée des ordinateurs personnels. Parmi ces films par ordinateur, *Arbitrary Logic* (1984-86)<sup>201</sup> constitue un travail exemplaire de traduction de données, en référence aux orgues à couleurs et à la musique visuelle (dont nous avons parlé au premier chapitre), qu'il cite abondamment dans les textes qu'il consacre au commentaire de ses travaux.

*Arbitrary Logic*, achevée en 1986, est une pièce générative et interactive<sup>202</sup> (par le biais de l'utilisation de la souris). Sur l'écran divisé en seize zones de surface différentes, des rectangles de couleur, d'échelle différente (de l'aplatissement à des bandes de pixels), évoluent tandis que des sons de synthèse sont joués simultanément par l'ordinateur, grâce à un synthétiseur MIDI<sup>203</sup>.

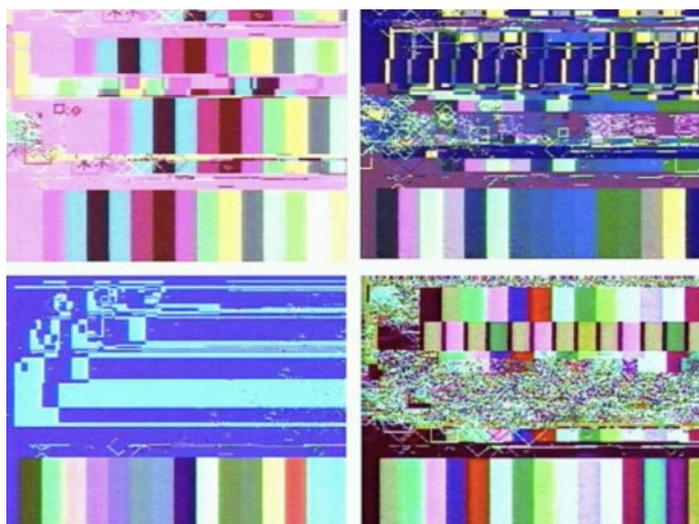


Fig. 35 : Malcolm Le Grice, *Arbitrary Logic* (1984-86)

Quatre captures d'écran

© LUX

201 Quelques captures vidéo de cette pièce interactive sont visibles dans ce documentaire en ligne, sur le site du LUX, à 5 min 48 s, mais sans le son, <https://www.youtube.com/watch?v=wCGIqRVRqmU>, consulté le 30/01/23.

202 Une interprétation de la pièce a été capturée et linéarisée par Le Grice en vidéo, puis intégrée à la suite de films *Sketches for a Sensual Philosophy* (1988), destinée à la télévision.

203 L'interface MIDI (Music Digital Interface), créée en 1983, permet de transmettre des informations musicales informatiques à des instruments compatibles. Ces informations peuvent être de nature musicale (le timbre par exemple) ou bien relative à un ordre d'exécution (une partition), et peuvent être déclenchées par un programme informatique.

Comme le souligne Bidhan Jacobs dans un ouvrage consacré à l'esthétique du signal, Malcolm le Grice « profite de la propriété de translatabilité (traduction et transfert) des données par l'ordinateur via le port MIDI qui peut générer, à partir des mêmes séquences de code, à la fois la visualisation de champs de couleurs et la sonification<sup>204</sup> ». La partie visuelle, composition semi-aléatoire multicolorée, s'apparente à du *Glitch* avant l'heure. Chaque zone de l'écran est associée à un timbre, un jeu interactif de vitesse et de choix est donc possible entre les seize timbres<sup>205</sup>. Le fondement de l'œuvre est l'utilisation d'un même programme qui génère à la fois l'image composite abstraite, et le son. Pour cette pièce, Malcolm le Grice a utilisé le langage Fast BASIC, sur un Atari ST<sup>206</sup>.

Dans *Arbitrary Logic*, les éléments visuels et sonores sont entièrement synthétisés. Aucune des données n'a été préalablement enregistrée, et le programme est entièrement conçu à partir de formules. Qui produisent les formes (rectangulaires) initiales, les agencements de couleur, dans des tonalités de rouge, de vert et de bleu, et la sélection des timbres instrumentaux. Le programme contrôle un système basique et cohérent de changement de valeurs, qui peut être modifié et interrompu à l'aide de la souris. Avec les boutons, on peut modifier les voix, les sons, le volume et la vitesse tout en contrôlant l'incursion de valeurs aléatoires. La souris contrôle également un système grâce auquel les rayures verticales sont copiées sur différentes parties de l'écran puis combinées (superposées) selon le système d'opérateurs logiques informatiques (et, ou, ni, etc.)<sup>207</sup>.

Le fait que ce dispositif puisse être utilisé sur scène<sup>208</sup>, comme « un instrument de musique dont il est possible d'apprendre à jouer, mais qui possède des structures de déve-

204 Bidhan Jacobs, *Esthétique du signal. Hacker le filmique*, Paris, Mimésis, coll. « Images, médiums », 2022, p. 423.

205 Philippe Langlois, « La musique des images », in Yann Beauvais (dir.), *Malcolm Le Grice. Le temps des images*, op. cit., p. 285.

206 Le ST fait référence à l'architecture 16/32 bits du microprocesseur. Sorti en 1985, l'Atari ST a eu un énorme succès, notamment avec les jeux vidéo, et a contribué à démocratiser la pratique de la musique par ordinateur, ainsi que le développement de la norme MIDI. L'un des ses concurrents directs était l'Amiga 500, de Commodore qui se démarquait notamment par une meilleure qualité de son.

207 Malcolm Le Grice, « L'art au pays de l'Hydre des médias », 1998, in Yann Beauvais, *Malcolm Le Grice. Le temps des images*, op. cit., p. 206-207.

208 La pièce a fait partie d'une performance *Live* de musique par ordinateur donnée avec le musicien Keith Rowe (qui pratique l'improvisation libre avec des guitares préparées) à la London Filmmakers Cooperative en décembre 1989.

loppement programmatiquement établies<sup>209</sup> », montre bien la relation dynamique entre l'interactivité et la générativité mises en jeu ici.

Il y a, dans le côté musical, une parenté avec un logiciel imaginé et développé par Laurie Spiegel, *Music Mouse* (1985-86)<sup>210</sup>. Celui-ci permettait, sur un ordinateur Apple (puis, plus tard, sur Amiga et Atari) de générer de la musique par le simple déplacement de la souris. En cochant ou décochant des paramètres, la nature de la musique pouvait évoluer en temps réel.

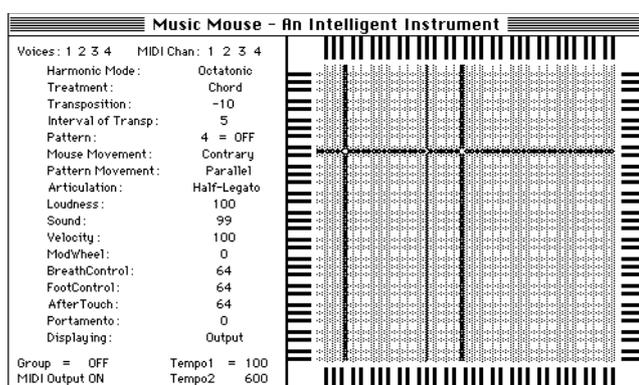


Fig. 36 : Laurie Spiegel, *Music Mouse* (1985-86)

Version pour Atari ST

© Laurie Spiegel, musicmouse.com

Dans *Arbitrary Logic*, l'aspect le plus important est qu'un même algorithme génère à la fois l'image et le son. L'artiste crée ses propres correspondances, « arbitraires », entre l'image et le son :

J'en suis arrivé à la conclusion qu'il n'y a pas de lien intrinsèque entre une couleur et un point particulier — d'où le terme « arbitraire » dans le titre — mais que la construction des relations entre son et couleur dans le « langage » abstrait de l'œuvre ainsi que l'expérience pour le spectateur restent artistiquement valables<sup>211</sup>.

209 Malcolm Le Grice, « Maintenant et alors. Réflexions et présence », 2003, in Yann Beauvais (dir.), *Malcolm Le Grice. Le temps des images, op. cit.*, p. 233.

210 Une documentation de *Music Mouse* est accessible sur le site de Laurie Spiegel, <http://musicmouse.com/>, consulté le 30/01/23.

211 Malcolm Le Grice, « Maintenant et alors. Réflexions et présence », 2003, in Yann Beauvais (dir.), *Malcolm Le Grice. Le temps des images, op. cit.*, p. 232.

Cela ne garantit toutefois en rien la bonne réception de ce procédé par le spectateur-auditeur. En effet, comme le souligne Philippe Langlois, celui-ci n'a pas une perception directe, mais intuitive de ces correspondances mises en place. Le rythme du son, et le mouvement des rayures en soi, n'est par exemple pas exploré sur le plan rythmique « [...] où les motifs, en lien avec la fréquence et le timbre, apparaissent la plupart du temps comme déconnectés<sup>212</sup> ». Dans un autre film informatique, *Digital Still Life*, Le Grice explore justement cet aspect en utilisant cette fois un système de génération de son calé sur le rythme et le séquençage d'images numérisées de natures mortes à basse résolution et retraitées.

### 1.3.3 L'aboutissement méconnu de la pensée de John Whitney : un outil de composition audio-graphique



Fig. 37 : John Whitney, *Moon Drum* (1991)  
© Moon Drum, Twelve evocations by John Whitney,  
VHS, Mystic Fire Video, 1991

Nous avons évoqué le travail de John Whitney à deux reprises, dans cette première partie dédiée à la musique visuelle programmée. Une première fois avec la série de films optiques *Five Film Exercices*, réalisée dans les années 1940 avec son frère James. Ils y inventent une forme inédite, à mi-chemin entre musique synthétique et animation visuelle : une musique audiovisuelle. La seconde évocation est relative aux films

<sup>212</sup> Philippe Langlois, « La musique des images », in Yann Beauvais (dir.), *Malcolm Le Grice. Le temps des images, op. cit.*, p. 285.

par ordinateur de John Whitney, seul. Si les formes visuelles, courbes mathématiques en mouvement, utilisent des systèmes d'harmonies empruntées au champ musical, une musique préexistante est « plaquée » sur l'animation, ce qui constitue une sorte de retour en arrière en ce qui concerne la relation audiovisuelle. D'une certaine manière, elle est plus pauvre que dans les films des années 1940.

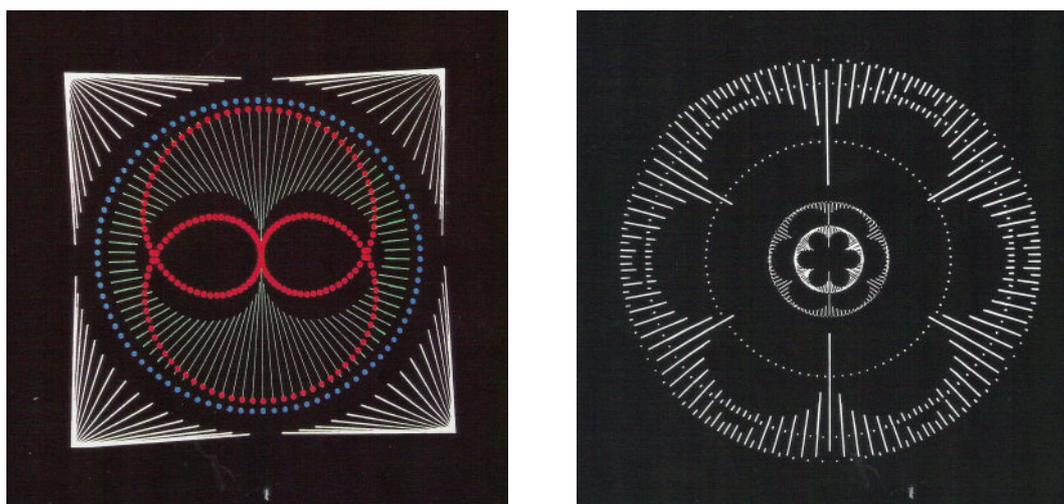


Fig. 38 : John Whitney, extraits de l'ouvrage *Digital Harmony. On the Complementarity of Music and Visual Art* (1980)

L'histoire des films de John Whitney s'arrête souvent, chez les historiens, à ces films informatiques des années 1960. Pourtant, l'objectif d'inventer de nouvelles formes audiovisuelles, que s'était fixé John Whitney dès le début de sa carrière, ne s'est pas arrêtée là. Au contraire, il publie en 1980 un ouvrage manifeste, *Digital Harmony*<sup>213</sup>, qui contient de nombreux articles inédits, historiques ou théoriques. Une partie d'entre eux, très didactiques, décrivent son parcours, ses inspirations, contiennent des analyses de ses travaux (on y trouve des exemples de code informatique utilisés pour générer certaines études). Par ailleurs, les articles importants déjà publiés s'y voient réunis. Il nous semble que cet ouvrage (même s'il s'avère très difficile à trouver) propose une somme d'informations indispensables à tous ceux qui s'intéressent aux rapports entre l'image et le son, au mouvement, ou bien au code de création.

Après la publication de cet ouvrage, et ce jusqu'à son décès en 1995, John Whitney se

213 John Whitney, *Digital Harmony*, op. cit.

concentre sur le développement d'un instrument logiciel qui permet de composer du visuel et du son en temps réel, en collaboration pour le développement avec Jerry Reed : le *Whitney-Reed RDTD* (1980-1995). À ce stade, John Whitney développe un usage pleinement artistique des outils informatiques, et a mis en œuvre bon nombre de concepts passionnants, une véritable science du mouvement en partie basée sur des réflexions déduites des structures musicales et de la composition sonore.

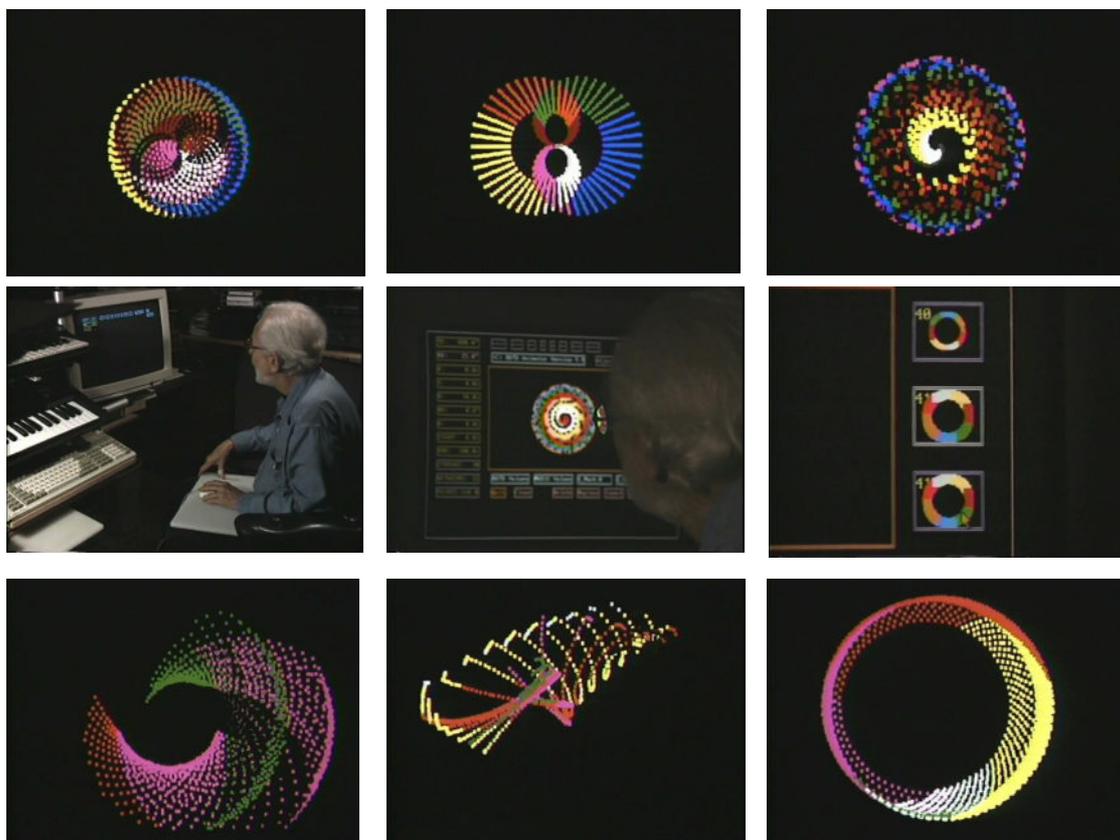


Fig. 39 : John Whitney utilisant le Whitney-Reed RDTD, avec des extraits de *Moon Drum* (1991)  
© The John Whitney Collection, DVD, Pyramid Media, 2009

Cette période est méconnue dans l'histoire de John Whitney, alors qu'elle synthétise finalement toutes les réflexions accumulées au cours de son parcours. Dans un documentaire précieux *A Personal Search : for the Complementary of Music and Visual Art* (1992), il nous présente son travail de l'époque à travers d'une part son inscription dans son parcours et d'autre part la présentation générale de son logiciel, avec quelques œuvres toutes récentes. Ce qui nous intéresse dans la première partie, c'est

qu'il va nous passer de longs extraits des *Five Abstract Film Exercices*, les seules œuvres qui selon nous, dans son parcours passé, présentent un véritable équilibre audiographique.

Dans la seconde partie, plusieurs études sont présentées. Cette fois-ci pour lui, le pixel est l'élément de base. Il peut, grâce à son outil, à la fois composer avec le son et le visuel, mais aussi le faire en temps réel. Si esthétiquement cela semble un peu désuet (probablement à cause du peu de couleurs disponibles), l'interaction entre le sonore et le visuel est particulièrement réussie et intrigante, tout comme l'étaient les *Five Abstract Film Exercices*. Il réalise à l'aide de ce logiciel *Spirals* (1987), puis *Moon Drum* (1991).

*Moon Drum* est probablement la plus méconnue de ses œuvres. Il s'agit d'un ensemble de 12 courts-métrages. Elle a été éditée en 1991 en cassette VHS chez Mystic Fire Video<sup>214</sup>, mais reste très peu projetée<sup>215</sup>. Les titres des 12 pièces sont : *Moon Drum* (qui donne son titre à la série), *Navajo*, *Hopi*, *Kwakiutl*, *Qxaquilt*, *Black Elk Requiem*, *Chaco*, *Mimbres Star*, *Chumash*, *Kachina*, *Chapala* and *Acoma*. Selon le texte de présentation présent sur la cassette, les 12 pièces sont une évocation impressionniste des croyances amérindiennes, selon lesquelles il existait des relations entre les couleurs et les tonalités musicales.

Il y a cinquante ans, dans l'obscurité et dans une poussière suffocante, je me suis rendu en voiture dans un endroit du Nouveau-Mexique pour observer des danses cérémonielles. C'était l'impact des tambours si forts qu'ils martèlent le cœur, ainsi que de brèves visions à la lumière aléatoire du feu, un serpent, les motifs et les symétries momentanées et les hallucinations d'une telle nuit, cela colore à jamais mes rêves et réflexions contenus dans ces douze pièces qui tentent d'évoquer l'ambiance, les couleurs, les sons et les œuvres d'art originaires de ce monde occidental [Nous traduisons]<sup>216</sup>.

214 Nous avons eu la chance d'en acquérir un exemplaire, que nous avons fait numériser.

215 Une version est disponible en ligne sur Internet Archive, <https://archive.org/details/JohnWhitneyMoonDrum1991>, consulté le 30/01/23. Elle nous paraît toutefois légèrement désynchronisée par rapport à notre version VHS numérisée.

216 « Fifty years ago, in darkness and in suffocating dust I drove to a location in New Mexico to observe ceremonial dances. It was the impact of drums so loud they pound in the heart, as well as brief visions in random firelight, a snake, the patterns and momentary symmetries and hallucinations of such a night, this forever colors my dreams and reflections contained in these twelve pieces which attempt to evoke the mood, the colors, sounds, and the artworks native to that Western world ».

Texte de présentation de la cassette VHS *Moon Drum*, Mystic Fire Video, 1991.

Dans *Moon Drum*, on retrouve une inspiration visuelle et sonore de ces rituels ancestraux, mais traités de manière nouvelle, dans un langage audio-graphique très poussé. En effet, il ne s'agit pas de mappages directs entre couleur et son, ou bien entre mouvement et musique, mais de véritable composition, en contrepoint, de l'image et du son. L'image et le son sont tous deux synthétiques, générés informatiquement. Cela a permis à Whitney de composer des va-et-vient entre la musique et les formes graphiques, qui se développent tantôt en synchronisation, tantôt de manière autonome, dans des mouvements de développement parfois proches de ceux de feux d'artifice discrets et lents. La perception du spectateur-auditeur suit donc ces développements comme s'il suivait les flux croisés de deux instruments ou de deux voix se succédant et se superposant.

La musique elle-même n'est pas un flux continu, mais est constituée d'un mélange de sons percussifs et de sons de durée variable, composée grâce à un clavier MIDI<sup>217</sup>. D'une certaine manière, le son seul, ou le visuel seul serait moins intéressant. C'est la relation image-son, ici qui est le cœur du travail. Cet équilibre architectural et perceptuel proposé par Whitney a été rendu possible par l'utilisation de son logiciel.

Lors d'un échange par mail avec lui, le développeur Jerry Reed nous a fourni un certain nombre d'informations sur le déroulement de sa collaboration avec John Whitney. Lorsqu'il le rencontre en 1986, l'idée est celle d'un partenariat à deux, en vue de créer un produit qui serait plus tard destiné à la vente. La version utilisée pour *Moon Drum* était encore un prototype. Au cours des 10 années suivantes, Reed a fait évoluer le logiciel par intermittence, en fonction des financements qui se présentaient à Whitney. Reed a fait en sorte que l'interface soit simple et directe à utiliser. Il a même développé une version avec une interface plus moderne, de style Windows, mais Whitney ne l'a jamais utilisée<sup>218</sup>.

---

217 Il semblerait que le clavier soit de la marque Kurzweil.

218 Nous avons pu échanger avec Jerry Reed par email. Vers 2005, celui-ci avait apparemment écrit un tout nouvel outil d'édition utilisant la plate-forme graphique de création d'applications Microsoft WPF (Windows Presentation Foundation), en y ajoutant des paramètres expérimentaux. Toutefois, cela n'a pas permis de restituer les profils originaux de couleurs et d'instruments. La version originale, quant à elle, ne peut fonctionner que sur le matériel original. Il est constitué d'une machine Compaq et d'un équipement MIDI spécialisé. L'un des trois fils de John Whitney, Michael, l'a conservé chez lui à Nashville. D'après un échange que nous avons eu avec son autre frère John Jr., l'ensemble est fonctionnel et les trois fils essaient de transférer numériquement, en pleine résolution, la série complète (l'équipement de John Whitney ne permettait pas à l'époque cette qualité).

Avec *Moon Drum*, nous sentons qu'émane de la partie visuelle une structure forte, mais qu'elle intègre aussi des éléments de « hasard orienté », et que le son synthétique fonctionne sur les mêmes principes. Nous sommes en présence de compositions bien plus complexes et conformes à ce que nous imaginons en termes d'« audiographique ».

Dans ces années, le travail des interfaces interactives et les possibilités croissantes de générativité ont mené les artistes, comme John Whitney, à développer leurs propres outils de manipulation d'image et de son, entre composition et instrument.

### 1.3.4 Golan Levin. Une suite d'instruments audiovisuels



Fig. 40 : Golan Levin, *Audiovisual Environment Suite* (1998-2000)  
© Golan Levin / flong.com

Dans les années 1980, avec l'avènement des micro-ordinateurs personnels, puis dans les années 1990, avec le réseau Internet, les formes audiovisuelles s'hybrident et se multiplient, de formes performatives aux installations, des applications téléchargeables par l'utilisateur, jusqu'aux formes en réseau. Dans ce contexte, les artistes commencent à utiliser des logiciels offrant des possibilités plus ouvertes de programmation, loin de la timeline linéaire des logiciels de montage habituels, ou programment — ou font programmer — des logiciels dédiés à leurs besoins. Par le numérique, il devient plus facile de manipuler l'image et le son grâce aux mêmes types de programme. Dans ce contexte de nouveaux types d'interfaces dédiées à la création au-

diovisuelle sont imaginées, à la frontière de l'animation abstraite et du jeu vidéo. Le temps réel, la générativité et l'interactivité qui deviennent envisageables par le numérique n'ont pas les mêmes objectifs. Le logiciel peut servir d'outil de composition non linéaire, mais pour un rendu audiovisuel fixé. Il peut servir d'outil de manipulation en temps réel, dans un contexte performatif. Il peut enfin être confié à des utilisateurs qui vont s'en servir librement.

### ***Interfaces a-musicales jouables***

Golan Levin, un artiste et chercheur, est peut-être celui qui a le mieux exploré cette histoire des logiciels d'art audiovisuel. Dans le cadre de son dernier cycle au MIT sous la direction de John Maeda<sup>219</sup>, il a dressé une histoire de ces logiciels et interfaces et il a lui-même développé une série d'applications-instruments dont les interfaces sont directement réactives aux gestes de l'utilisateur.

Dans son article « Audiovisual Software Art », publié dans le catalogue de l'exposition *See This Sound*<sup>220</sup>, il dresse un corpus et classifie ces dispositifs selon différents rapports entre image et son qui y sont mis en jeu sous différents angles : la visualisation et la notation, la transmutabilité, la performativité et la générativité. Ces relations peuvent s'entremêler dans une même œuvre.

Certaines œuvres génèrent des images ou des animations à partir de sons en direct ou préenregistrés. D'autres projets génèrent de la musique ou des sons à partir d'images statiques ou de signaux vidéo, ou utilisent des interfaces graphiques sur écran pour contrôler de manière interactive les processus musicaux. D'autres œuvres d'art génèrent à la fois du son et des images à partir d'une source non audiovisuelle d'informations externes (par exemple des données boursières, des données de capture de mouvement humain) ou à partir d'un processus aléatoire interne. [Nous traduisons]<sup>221</sup>.

---

219 John Maeda est l'un plus grands contributeurs du design interactif et est à l'origine du terme « code de création ». Il a dirigé pendant 8 ans l'Aesthetics and Computation Group (ACG), au sein du MIT Media Lab à Cambridge. Parmi ses travaux, les *Reactive Books* (1993-1999), dont l'un est réactif au son. Le logiciel Processing, de Casey Reas et de Ben Fry, est une évolution d'un logiciel de Maeda, *Design by Numbers*. Celui-ci décrit dans l'ouvrage éponyme *Design by Numbers* (1999). Le travail de Maeda avec ses étudiants est documenté dans l'ouvrage *Code de création* (2004). Il a également publié l'ouvrage *De la Simplicité* (2007), dans lequel il prône le recours à un design épuré et minimaliste, fondé sur la fonctionnalité.

220 Nous avons évoqué à plusieurs reprises cette exposition importante, notamment dans l'introduction.

221 Golan Levin, « Audiovisual Software Art », in Dieter Daniels et Sandra Naumann, *op. cit.*, p. 271.

On peut ajouter à ce corpus des formes plus récentes, qu'Hervé Zénouda nomme les « audio-games », c'est-à-dire des applications disponibles sur consoles, ordinateurs ou téléphones portables, à la frontière du jeu et de la création musicale, « dans un espace situé entre l'instrument, la boîte à musique et les automates musicaux ». En référence à un article d'Alain Bonardi et d'Alain Rousseau, il qualifie ce type d'interfaces d'a-musicologiques, c'est-à-dire « utilisant des symboles ne faisant pas référence à la musicologie classique pour manipuler et produire du sonore<sup>222</sup> ». Zénouda voit dans cette nature d'œuvre une filiation entre trois courants : la musique audiovisuelle de John et James Whitney, les partitions graphiques des musiques expérimentales et contemporaines, et les interfaces graphiques pour la composition musicale.

### ***Une substance audiovisuelle***

Il est impossible de citer ici toutes ces applications-instruments<sup>223</sup>, mais Golan Levin, dans le contexte de sa recherche au MIT, a développé lui-même entre 1998 et 2000 une série d'instruments audiovisuels. Il l'a nommée *AVES* (Audiovisual Environment Suite)<sup>224</sup> et la décrit ainsi :

*AVES* est une nouvelle métaphore d'interface informatique pour la performance simultanée et en temps réel d'images et de sons dynamiques. Cette métaphore découle de l'idée d'une « substance » audiovisuelle, temporelle, infiniment variable et inépuisable qui peut être créée par gestes, déposée, manipulée et effacée dans un espace imagé à structure libre et sans schéma. La métaphore de l'interface est illustrée par cinq systèmes de synthèse interactive et audiovisuelle, dont les dimensions sonores et visuelles sont très plastiques, malléables

---

222 Hervé Zénouda, « Nouvelles organologies du sonore : les audio-games. La question des interfaces a-musicologiques », *Interfaces numériques* Vol. 1, n° 3, 2012, <https://www.unilim.fr/interfaces-numeriques/2456>, consulté le 30/01/23.

223 Parmi elles, *Music Insects* (1991-2004), de Toshio Iwai (également créateur du jeu *Electroplankton* en 2005), permettait de générer des sons à l'aide de petits insectes abstraits, pixelisés, qui venaient se cogner contre des pixels sonores. En ajoutant ou retirant des pixels, l'utilisateur créait des chemins pour les insectes, et donc sa propre musique. Chaque insecte était associé à un son différent. Autre exemple, les applications de Brian Eno et de Peter Chilvers, comme *Bloom* (2008), qui produisent simultanément de la musique électronique *Ambient* et génèrent des formes graphiques primaires en fonction des actions de l'utilisateur. Globalement, il s'agit souvent d'un mappage entre des positions de souris ou de forme, et des paramètres sonores. Le site *Soundtoys*, recense une grande partie de ces expériences multimédia. Beaucoup d'entre elles sont malheureusement obsolètes car programmées à l'époque sous Flash. Le site est accessible à l'adresse <http://www.soundtoys.net/>, consulté le 30/01/23.

224 Une version est disponible en ligne sur le site du MIT, <https://acg.media.mit.edu/people/golan/aves/>, consulté le 30/01/23.

et étroitement liées par des topographies (mappings) déterminées par les perceptions<sup>225</sup>.

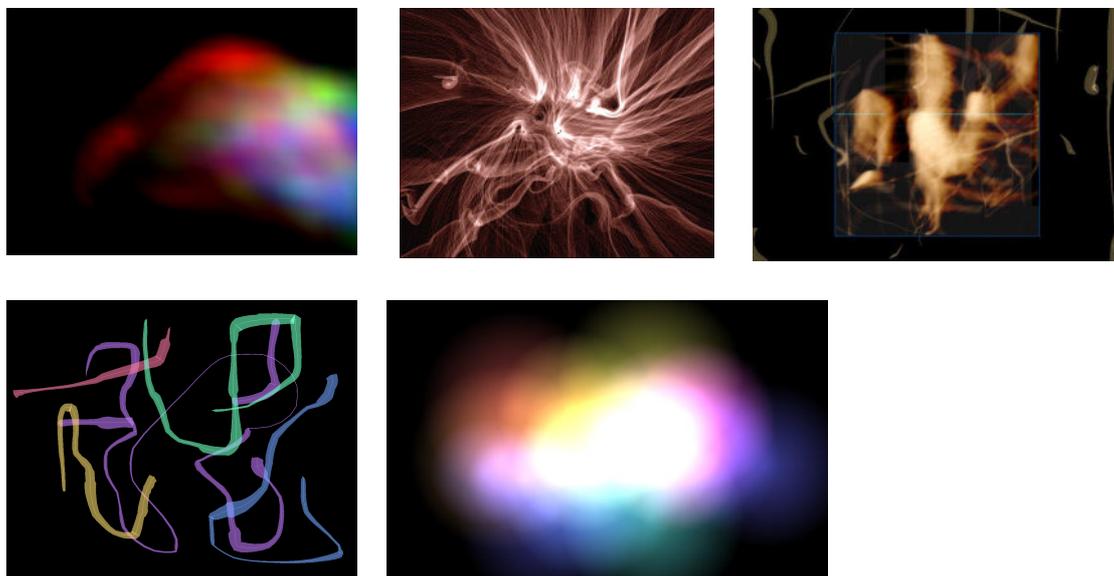


Fig. 41 : Golan Levin, *Audiovisual Environment Suite* (1998-2000)  
de gauche à droite : *Aurora* (1999), *Floo* (1999), *Yellowtail* (1999), *Warbo* (2000) et *Loom* (1999)  
© Golan Levin/The aesthetics + computation group

On retrouve la notion de « substance » qu'utilise également Michel Fano et évoquée à propos du travail de Lillian Schwartz au deuxième chapitre. Les 5 modules principaux sont *Aurora* (1999), *Floo* (1999), *Yellowtail* (1999), *Warbo* (2000) et *Loom* (1999). La plupart de ces applications, comme *Aurora*, *Yellowtail* et *Loom* fonctionnent sur le principe de la récupération des informations d'un tracé (à la souris ou au stylo sur tablette numérique) ou d'un mouvement de souris. Les dessins créés sont ensuite redessinés, et selon les modes de fonctionnement, génèrent du son grâce à des techniques de synthèse du son (synthèse additive, par modulation de fréquences ou synthèse granulaire). Pour Hervé Zénouda, il s'agit d'une forme de nouvelle lutherie audiovisuelle<sup>226</sup>. Le module *Aurora* donne la possibilité à l'utilisateur de créer une nébuleuse de couleurs et de sons (très proche des formes Lumia de Thomas Wilfred dans les années 1920), qui réagit aux mouvements de la souris. *Floo* permet la fabrication d'une composition filaire réactive et la génération de drones sonores granulaires synchroni-

<sup>225</sup> Biographie de Golan Levin sur le site de la Fondation Langlois.

<https://www.fondation-langlois.org/html/f/page.php?NumPage=230>, consulté le 30/01/23.

<sup>226</sup> Hervé Zénouda, *Les images et les sons dans les hypermédias artistiques contemporains*, Paris, L'Harmattan, 2004, p. 195.

sés. *Yellowtail* autorise le dessin de lignes ; quand celle-ci se retrouve dans une zone rectangulaire particulière, visible, des paramètres liés à leur qualité de forme et à leur épaisseur sont sonifiés<sup>227</sup> grâce à une synthèse additive<sup>228</sup>. Dans *Warbo*, des gouttes lumineuses (assez similaires à celles d'*Aurora*) sont générées. Lorsque l'on passe dessus, des accords de son sont produits. *Loom*, enfin, est un dispositif de tracés qui se redessinent et qui sont lus et convertis en son en temps réel, selon leur courbure, par un synthétiseur FM<sup>229</sup>.

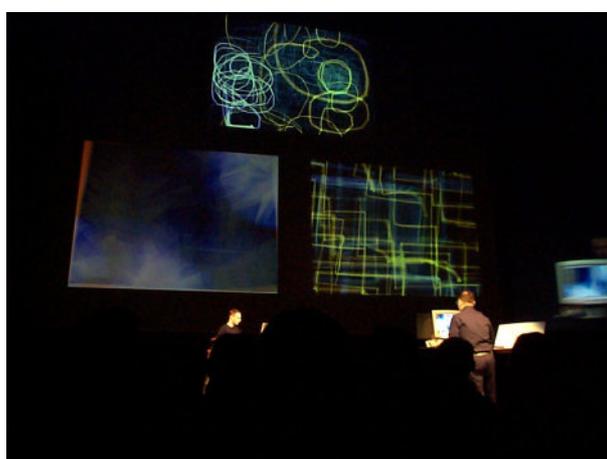


Fig. 42 : Golan Levin, *Scribble* (2000), performance audiovisuelle  
© Golan Levin/The aesthetics + computation group

En 2000, Golan Levin crée la performance audiovisuelle, *Scribble*<sup>230</sup>, à Ars Electronica. Celle-ci est basée sur l'utilisation des modules d'AVES. En duo ou trio, les manipulations sont effectuées en direct et projetées sur grands écrans et système de diffusion sonore<sup>231</sup>. Si ces outils sont intéressants, et que leurs possibilités expressives sont étendues, le résultat reste assez peu architecturé, au sens où les langages mis en

227 C'est-à-dire convertis en son selon des règles de mappage. Nous reviendrons sur ces principes dans le dernier chapitre (1.4) de cette partie.

228 « La synthèse additive est une classe de techniques de synthèse sonore basée sur l'addition de formes d'onde élémentaires pour créer une forme d'ondes plus complexe » (Curtis Road, *L'audionumérique. Musique et informatique*, Paris, Dunod, coll. « Audio-Photo-Vidéo », 2e édition, 2007, p. 387).

229 Synthétiseur basé sur le processus de modulation de fréquence, inventé par John Chowning dans les années 1970.

230 Une captation est disponible en ligne sur le site de Golan Levin, <http://flong.com/archive/projects/scribble/index.html>, consulté le 30/01/23.

231 Quelques années plus tard, en 2003, Cécile Babiolo, s'inspirant des ardoises magiques, dessinait suivant les mêmes types de processus, dans sa performance *Scribbling Waves* les sons en 3D et en temps réel : « le même mouvement de la main module le son et trace les lignes que l'on voit simultanément se former sur l'écran » (Cécile Babiolo, <http://babiole.net/scribbling-waves/>, consulté le 30/01/23).

œuvre ne semblent pas poser suffisamment de contraintes : cette liberté trop grande donne une sensation de flottement. Cette sensation est renforcée par le fait que les modules ne sont a priori pas pensés au départ pour être mis ensemble.

### 1.3.5 L'exploration de la relation audiovisuelle dans les formes semi-autonomes d'Antoine Schmitt

Le processus génératif est l'un des fondements des œuvres évoquées précédemment. Selon Geoff Cox, Alex McLean et Adrian Ward, « les médias génératifs reposent sur le principe fondamental d'une modification des données à mesure que le code s'exécute<sup>232</sup> ». Philip Galanter, artiste et théoricien, définit quant à lui, dans son article « What is Generative Art ? » l'art génératif comme une « une pratique où l'artiste crée un procédé, par exemple un ensemble de règles langagières, un programme informatique, une machine ou tout autre mécanisme qui est par la suite mis en marche et qui, avec un certain degré d'autonomie, entraîne la création d'une œuvre d'art issue de ce procédé<sup>233</sup> ». Il pointe trois *clusters* contemporains dans lesquels s'exerce l'art génératif : la musique électronique (au sens numérique) et algorithmique, l'infographie et l'animation, la scène démo et la culture VJ. Le travail d'Antoine Schmitt, qui explore la question des objets à comportements, s'inscrit au croisement de ces trois mouvances.

#### *Les nanoensembles, une Gestalt audiovisuelle.*

En 2002, Antoine Schmitt crée sa *nanomachine*<sup>234</sup>, un ensemble de petits modules audiovisuels (les *nanoensembles*) qui peuvent avoir une forme linéaire (installation ou animation web) ou bien une forme performance. Dans ces pièces, Schmitt fabrique à la fois l'image et le son.

232 Geoff Cox, Alex McLean, Adrian Ward, « L'esthétique du code génératif » (The Aesthetics of Generative Code), 2001, a été traduit dans le cadre de l'ouvrage *ART ++* publié aux Éditions HXX, en juin 2011, sous la direction de David-Olivier Lartigaud. Il est uniquement disponible en ligne. [https://monoskop.org/images/a/ae/Cox\\_McLean\\_Ward\\_2000\\_2011\\_L\\_esthetique\\_du\\_code\\_generatif.pdf](https://monoskop.org/images/a/ae/Cox_McLean_Ward_2000_2011_L_esthetique_du_code_generatif.pdf), consulté le 30/01/23.

233 Philip Galanter, « What Is Generative Art? », 2003, [http://philipgalanter.com/downloads/ga2003\\_what\\_is\\_genart.pdf](http://philipgalanter.com/downloads/ga2003_what_is_genart.pdf), consulté le 30/01/23.

234 Une captation est disponible en ligne le site d'Antoine Schmitt, <https://www.antoineschmitt.com/nanomachine-fr/>, consulté le 30/01/23.

Il va, pendant sa performance, piocher dans une bibliothèque des objets graphiques primaires, qu'il va positionner à l'écran, dimensionner, puis activer. La bibliothèque, discrète, est visible, placée en bas de l'écran. Le spectateur peut ainsi assister aux manipulations de l'artiste et comprendre de manière intuitive quels sont les actions et objets en jeu.

Les sons (échantillons en boucle), les mouvements, et les comportements des objets primaires sont prédéfinis par programmation, ce sont les interactions que Schmitt va provoquer en manipulant les objets à l'écran qui vont déterminer les évolutions (chocs, boucles, rebonds, etc.) du *nanoensemble* qui se construit progressivement. Chaque objet primaire possède en effet un comportement qui lui dicte comment il doit se comporter si par exemple il rentre en collision avec un autre objet. L'artiste injecte également dans le système de légères instabilités qui produisent des variations légères, mais dérangeantes.

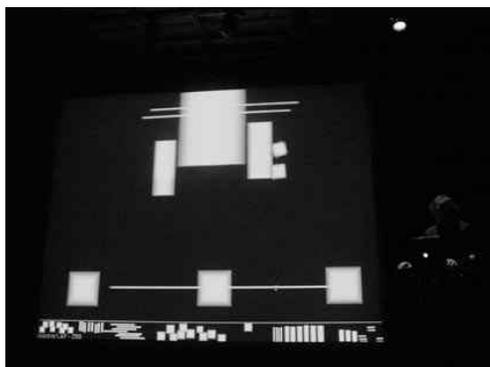


Fig. 43 : Antoine Schmitt, *Nanomachine* (2002-2004)  
Performance audiovisuelle  
© nouveauxmedias.net

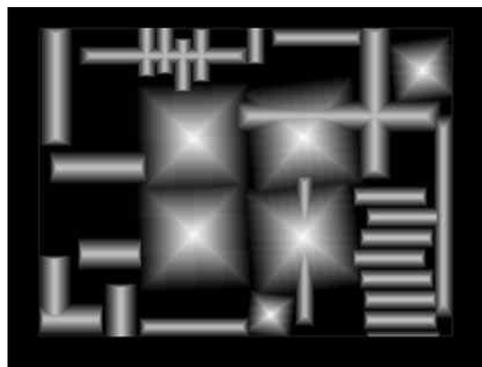


Fig. 44 : Antoine Schmitt, *Nanoensembles* (2002)  
Systèmes autonomes en ligne  
© nouveauxmedias.net

Dans le programme d'Antoine Schmitt, il n'y a ni maître ni esclave, pas de son qui dirige l'image et pas d'image qui dirige le son, mais un ensemble de micro-objets audiovisuels qui s'articulent pour former un ensemble stratifié. Au niveau perceptif, qu'il y ait peu ou beaucoup d'objets à l'écran, il y a une sensation d'équilibre entre l'image et le son.

Un autre aspect important est l'usage du noir et du blanc, associé à des questions essentielles dans les arts visuels et les arts graphiques. Il renvoie, par exemple, aux premiers films abstraits de Hans Richter et de Viking Eggeling, aux études d'Oskar Fischinger, aux films structuraux de Peter Kubelka, à l'art optique, au travail du graphiste Armin Hofmann, ou bien encore aux pionniers de l'art par ordinateur. L'un d'entre eux, Manfred Mohr, a travaillé pendant trente années exclusivement en noir et blanc, même s'il était tout à fait possible d'intégrer d'autres paramètres visuels.

J'appelle mon travail « Génératif » parce que toutes mes œuvres sont générées suivant des algorithmes (procédures logiques) établis par moi-même à l'avance. C'est là que réside fondamentalement ma contribution à la recherche esthétique. Je crée des signes, des « ÊTRES GRAPHIQUES » à partir d'un contexte rationnel. Ces signes ne renvoient qu'à eux-mêmes, et leur contenu témoigne de leur fabrication. [...] Mon travail est essentiellement linéaire. L'utilisation exclusive du blanc et du noir comme composants visuels permet de décrire un système rigoureux de décisions binaires<sup>235</sup>.

Le compositeur Iannis Xenakis, dans son ouvrage *Musiques Formelles* (1963), décrivant les phases fondamentales d'une œuvre musicale, évoquait la définition « d'êtres sonores et de leur symbolique communicable dans la mesure du possible (sons d'instruments de musique, sons électroniques, bruits, ensemble d'éléments sonores ordonnés, constitutions granulaires ou continues, etc.)<sup>236</sup> ».

Des questions de Gestalt<sup>237</sup> graphique et sonore sont aussi également en jeu. Quand peu d'objets audiovisuels sont à l'écran, on associe facilement une forme animée à son d'origine. Il y a une sorte d'accrochage, d'amorce, que le cerveau arrive à maintenir. En revanche, lorsque l'on dépasse un certain nombre d'éléments, le cerveau n'est plus capable de distinguer les évolutions de tous ces objets audio-visuels, et cherche à appréhender de plus grands ensembles. La même chose se produit pour le son. Le spectateur auditeur recrée alors de nouveaux rapprochements qui lui sont personnels, selon sa perception propre.

---

235 Manfred Mohr, cité dans Abraham A. Moles, *Art et ordinateur*, *op. cit.*, p. 308.

236 Iannis Xenakis, *Musiques formelles*, *op. cit.*, p. 33.

237 Psychologie de la forme. Nous y reviendrons dans la quatrième partie, section 4.1.5.

### ***Display Pixel 3, performance avec Vincent Epplay***

Dans une autre de ses performances audiovisuelles, *Display Pixel*, en 2003, Antoine Schmitt collabore avec le musicien Vincent Epplay. Dans cette troisième version, Antoine Schmitt a préparé des « scènes graphiques » que Vincent Epplay va déclencher selon les sons qu'il va choisir de jouer. Le système de Schmitt est par ailleurs capable de capturer les sons joués par le musicien, pour ensuite les cloner, ou les modifier.

Une fois les scènes graphiques à l'écran, Antoine Schmitt module leur comportement en temps réel. Il y a parfois des liens préalables entre l'image et le son, qui vont se défaire progressivement en fonction des manipulations. Hervé Zénouda décrit ainsi les systèmes de Schmitt en relation avec les fragments sonores d'Epplay :

Le musicien voit les réactions du système à l'écran et module son jeu en conséquence. Les pistes de la table de mixage lui permettent de sélectionner les parties de la musique qui entrent dans le programme graphique. Le musicien a donc le loisir de décider à tout moment des pistes qu'il injecte dans le système. Il peut tout aussi bien couper entièrement l'entrée du signal sonore, improvisant sur le jeu autonome du tableau graphique, jeu s'appliquant à des échantillons sélectionnés quelques instants plus tôt. Le plasticien/programmeur peut à tout moment influencer sur les tableaux interactifs en manipulant certains paramètres comme le tempo ou le nombre d'éléments graphiques présents à l'écran<sup>238</sup>.

La performance est constituée de 10 séquences interactives qui se succèdent, dans lesquelles les modalités d'interaction entre image et son diffèrent. Il y a toutefois globalement un principe général : un système génératif graphique, avec un comportement préétabli, est mis en route, nourri par des sons du musicien. Le plasticien et le musicien influent chacun quand ils le souhaitent (une forme d'improvisation) sur le système, à la manière d'un document partagé modifié en direct par deux éditeurs.

### ***Tempest. Flux de micro particules***

Dans une performance plus récente, *Tempest* (2012)<sup>239</sup>, Schmitt va collaborer avec l'artiste sonore Franck Vigroux, de manière plus déliée. Schmitt a cette fois travaillé sur la

<sup>238</sup> Hervé Zénouda, *op. cit.*, p. 177.

<sup>239</sup> Une captation est disponible en ligne sur le YouTube de la Compagnie d'Autres Cordes, <https://www.youtube.com/watch?v=0z9zYpMTFnU>, consulté le 30/01/23.

programmation de flux de micro-particules blanches (sur fond noir), grâce à divers algorithmes informatiques permettant de modeler des formes générales, comme des essaims d'insectes ou des tourbillons en mouvement permanent.

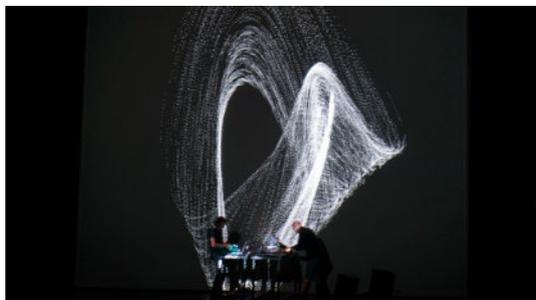


Fig. 45 : Antoine Schmitt et Franck Vigroux  
*Tempest* @ MUTEK Montreal 2013  
© Hayeur



Fig. 46 : Antoine Schmitt et Franck Vigroux  
*Tempest* (2012)  
© Galerie Charlot

Franck Vigroux travaille donc en direct le son, de manière plus analogique, tandis que Schmitt manipule ses flux graphiques, lui aussi en direct. Les deux artistes sont face à face, perpendiculaires à la scène, symbolisant ainsi l'équilibre entre l'image et le son.

Au final, le lien entre l'image et le son résulte entièrement du dialogue entre les deux performers que nous sommes, à l'oreille et au regard, avec quelques rendez-vous temporels. Au départ, nous avons prévu une possible liaison physique entre les deux instruments, sonores et visuels. C'est quelque chose que j'ai énormément expérimenté dans le passé, à la fois en performance et en installations. Quelque chose que j'ai même épuisé. Dans *Display Pixel*, performance avec Vincent Epplay de 2003, nous avons balayé l'ensemble des possibilités de rapport causal entre le son, l'image et le processus. Toutes les combinaisons ont été explorées et elles constituaient même l'essence du spectacle. Dans la *Nanomachine* (2002), un programme que je manipulais fabriquait en live à la fois le son et l'image, en de complexes machines folles construites devant le regard du spectateur<sup>240</sup>.

Il n'y a pas ici d'appairages préalables entre l'image et le son (si ce ne sont des répétitions et autres calages), la place est laissée libre à l'improvisation et à l'intuition.

<sup>240</sup> Antoine Schmitt, « *Tempest* : cosmogonie d'un spectacle », <https://www.antoineschmitt.com/tempest-cosmogonie-dun-spectacle/>, consulté le 30/01/23.

### 1.3.6 Générativité archéologique. Robert Henke et le CBM 8032

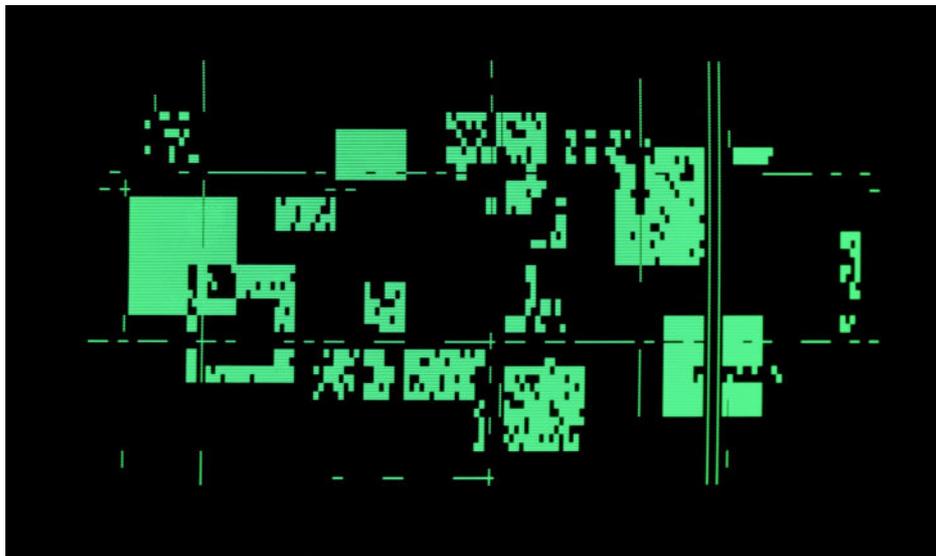


Fig. 47 : Robert Henke, *CBM 8032 AV*  
Performance audiovisuelle (2016-2021)  
© Robert Henke

De 2016 à 2021, l'artiste-musicien Robert Henke a travaillé sur une performance audiovisuelle dont l'objectif était de s'appuyer uniquement sur l'utilisation d'un ordinateur du début des années 1980<sup>241</sup>. Il choisit le Commodore CBM 8032, modèle symbolique du passage à l'ère informatique, pour développer le projet. Le titre de la performance, *CBM 8032 AV*, est directement tiré de ce nom. L'ordinateur étant très limité au niveau sonore (pas de carte son) et graphiquement (25 lignes de 80 caractères sur un écran monochrome vert), et pour permettre plus de combinaisons possibles, il crée un réseau constitué de cinq de ces ordinateurs, restaurés. Trois d'entre eux seront dédiés à la génération de sons, un à la génération graphique, le dernier au contrôle des quatre autres.

Lors de la performance, seules sont opérées des manipulations sur les ordinateurs, avec simplement en complément l'aide d'un rack d'effets audio, datant lui aussi du début des années 1980. Devant la projection sur grand écran, les moniteurs des ordinateurs sont tournés vers le public, afin qu'il puisse voir les interfaces et les manipulations de l'artiste, qui va interagir avec l'image et le son en direct.

<sup>241</sup> Une captation est disponible en ligne sur le site de FACT Magazine, <https://www.youtube.com/watch?v=CY1RAs7x74w>, consulté le 30/01/23.

Au début de la performance, l'écran de projection montre deux colonnes de données. Celle de gauche contient des séries de codes hexadécimaux, celle de droite présente des lignes constituées de pixels et de petits rectangles immobiles. Dans cet espace, des chiffres commencent à changer, le public comprend que c'est Henke qui est en train de saisir des informations au clavier. Il déclenche alors le premier processus de séquençage (c'est-à-dire une série de commandes qui pilotent des séquences temporelles) de la performance. Dans la colonne de droite, des curseurs commencent à s'animer dans un mouvement horizontal de gauche à droite, à passer sur certains fragments de nombres. Dans la colonne de gauche, au rythme du curseur, défilent les informations numériques qui semblent être liées au son.



Fig. 48 : Robert Henke, *CBM 8032 AV* (2016-2021)  
Début de la performance audiovisuelle  
© Robert Henke

Grâce à cette amorce, Henke introduit en quelque sorte la performance. Il nous fait comprendre qu'un langage de programmation de bas niveau (la saisie de code hexadécimal) a été utilisé, et que le seul mode d'interaction d'Henke avec l'ordinateur réside dans cette saisie au clavier d'instructions codées. Il s'agit, en résumé, d'une introduc-

tion sous forme de *Live Coding*, une pratique courante chez les artistes-programmeurs<sup>242</sup>.

Après ce court lancement, tout disparaît, et la performance en elle-même débute. On assiste alors à des enchaînements de séquences entremêlant formes graphiques vertes sur fond noir et sons expérimentaux.

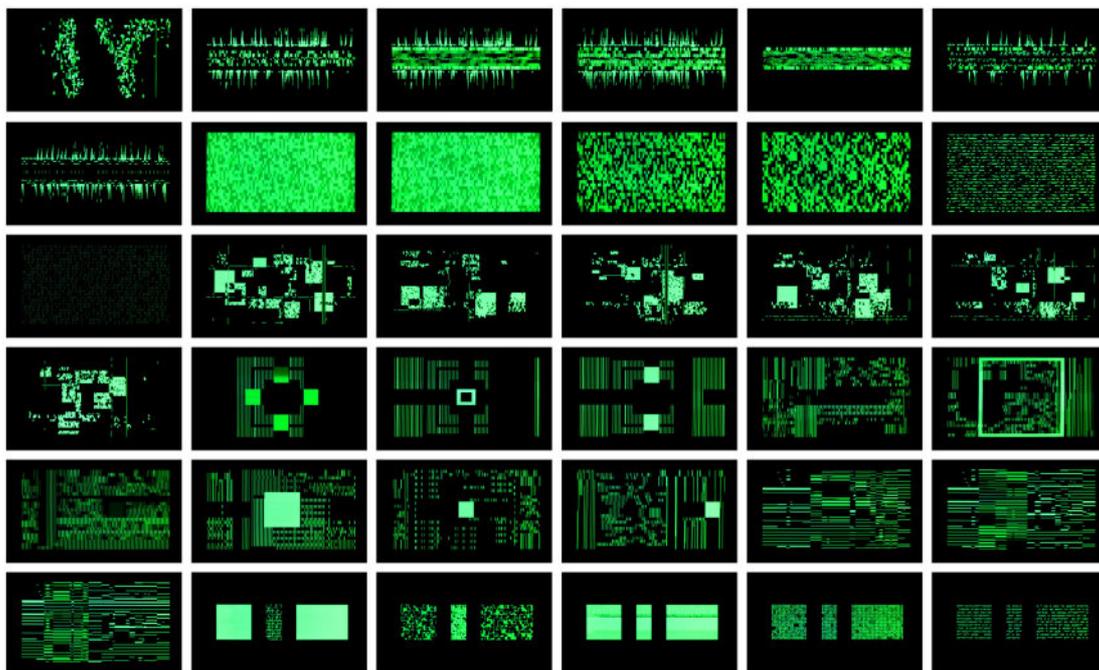


Fig. 49 : Robert Henke, *CBM 8032 AV* (2016-2021)

Captures de moments

© Robert Henke

Toutes ces images, contraintes par le système graphique de l'ordinateur, sont construites à partir de combinaisons issues des 256 lettres et symboles graphiques du jeu de caractères PETSCII de Commodore (les caractères graphiques étant privilégiés).

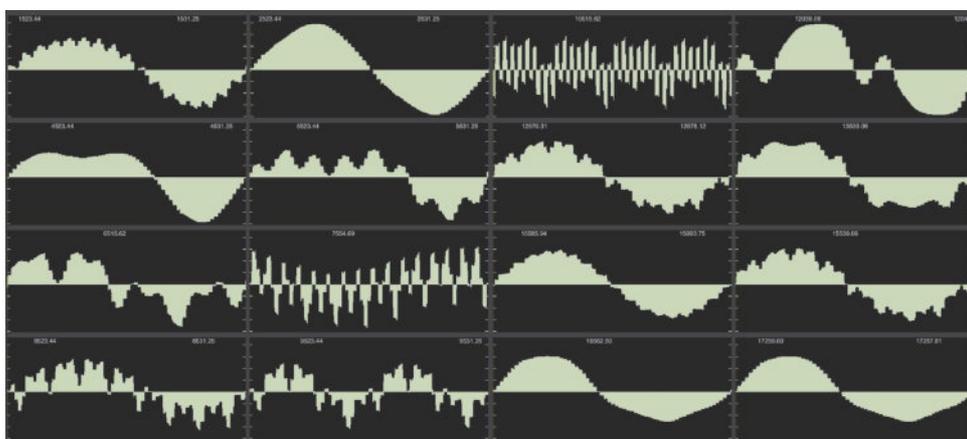
<sup>242</sup> Pendant une performance, la projection (généralement un écran divisé en deux parties) permet de voir l'artiste changer des lignes de code sur son interface de programmation, tout en voyant en direct le résultat.

Fig. 50 : Robert Henke, *CBM 8032 AV* (2016-2021)

Jeu de caractères PETSCII

© Robert Henke

Les sons très minimalistes sont assez similaires à ceux entendus chez Ryoji Ikeda<sup>243</sup> : bruits numériques, clics, ondes sinusoïdales, bips, etc. Henke explique que pour le son, il a dû utiliser un circuit imprimé fabriqué spécialement (une sorte de DAC), pour convertir la sortie 1 bit de l'ordinateur en 8 bits. Il s'agit du seul véritable hack, qui permet d'étendre les possibilités sonores.

Fig. 51 : Robert Henke, *CBM 8032 AV* (2016-2021)

Vocabulaire sonore. Planche de signaux utilisés

© Robert Henke

Pour programmer ce processus de génération de son et d'image, Henke a dû apprendre le « langage machine » de l'ordinateur. Celui-ci permet de coder directement une puce, grâce à des instructions numériques très abstraites. Il a ainsi réussi, avec l'aide d'une

<sup>243</sup> Artiste dont nous parlerons au prochain chapitre, à propos de l'utilisation des données.

équipe, à programmer des « routines<sup>244</sup> » capables de contrôler les différentes cartes de l'ordinateur.

La réalisation du projet 8032 AV comprenait de nombreuses étapes, cela m'a obligé à apprendre le langage machine 6502, à comprendre la conception du circuit imprimé du 8032, à me mettre dans la tête des personnes qui ont inventé cette machine. Il a fallu construire du nouveau matériel autour de lui. Et cela nécessitait de penser la génération du son et du graphisme dans des termes bien plus abstraits qu'avec n'importe quel outil contemporain [Nous traduisons]<sup>245</sup>.

Une fois les possibilités de programmation opérationnelles, Henke a créé un vocabulaire graphique et sonore, contraint par les limitations de la machine. Ce qui pourrait constituer une frustration est ici décrit par Henke comme un challenge, celui d'arriver à créer un langage audio-visuel avec des moyens réduits. Il se réfère à ce propos à l'esthétique programmée minimaliste du pionnier de l'art par ordinateur Manfred Mohr (dont nous avons parlé à plusieurs reprises).

### ***Incisions programmationnelles sur l'objet fermé***



Fig. 52 : Cory Arcangel  
*Super Mario Clouds* (2002)  
© Cory Arcangel

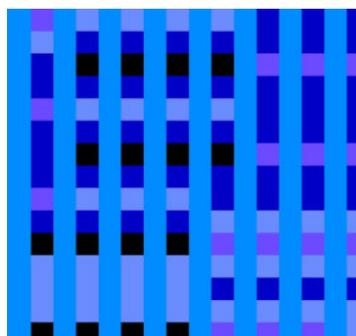


Fig. 53 : Cory Arcangel & Paper Rad  
*Super Mario Movie* (2005)  
© Cory Arcangel / coryarcangel.com

La démarche de reprogrammation d'Henke pourrait dialoguer avec une œuvre emblématique de l'artiste Cory Arcangel, *Super Mario Clouds* (2002). Il s'agit d'un jeu vidéo

244 En informatique, des fragments de code dédiés à l'exécution de tâches particulières, comme ici le contrôle de l'interface graphique et la génération de sons.

245 Page de présentation du projet, <https://roberthenke.com/concerts/cbm8032av.html>, consulté le 30/01/23.

modifié (un *Mod*) dont la source est le fameux jeu *Super Mario Bros* disponible à l'époque sur la console Nintendo NES. Arcangel a démonté la cartouche pour en retirer la puce et la remplacer par une autre qu'il a reprogrammée pour faire disparaître du jeu tous les éléments graphiques, excepté les nuages. Ne subsistent donc qu'un travelling sur un ciel bleu et des nuages. Pour le présenter sur le web, l'artiste a utilisé le format d'image *GIF*, qui autorise l'animation image par image (la vidéo à ce moment n'était pas possible, compte tenu des formats et de la bande passante disponible). En complément, comme pour la plupart de ses projets, Arcangel propose un mode d'emploi permettant de réaliser la même chose « à la maison ». La cartouche, objet matériel fermé, est ici ouverte par Arcangel dans une intervention de programmation. Dans un texte intitulé « Bricodage », Gilles Rouffineau voit dans *Super Mario Clouds* une tentative d'ouverture de l'objet technique théorisé par Simondon, théoricien de la technique.

La technologie — dont l'informatique — opère dans un domaine qui n'est plus accessible à la main humaine. Il n'est pas possible de « bricoler » physiquement à l'intérieur d'une puce électronique. En revanche, on peut en changer la destination en la (re) programmant. La programmation est le moyen de « reprendre la main » sur les « objets clos », un concept de Gilbert Simondon [...]<sup>246</sup>.

Le GIF n'autorisant pas le son, l'animation *Super Mario Clouds* est muette. L'extension du principe de hacking des puces de la cartouche du jeu à la musique est explorée en 2005 par Arcangel, en collaboration avec le collectif artistique Paper Rad dans l'animation *Super Mario Movie* (2005). Il s'agit toujours d'un *Mod*, mais cette fois-ci du jeu vidéo *Super Mario Brothers*, toujours sur Nintendo. Comme l'explique Arcangel<sup>247</sup>, il a lui-même appris à programmer le moteur d'animation Nintendo, puis il a transmis cet apprentissage aux membres de Paper Rad. La bande-son, conçue sur le même principe que la partie visuelle, a été composée comme celle d'un film par Jacob Ciocci de Paper Rad. Il utilise à la fois des morceaux de mélodies du jeu original, qui reviennent régulièrement, mais entremêlés à des fragments nouveaux. La puce sonore de la Nintendo

---

246 Gilles Rouffineau, « Bricodage » in David-Olivier Lartigaud (dir.), *Art++*, op. cit., p. 323.

247 Conférence de Cory Arcangel au Smithsonian American Art Museum (2009), YouTube, <https://www.youtube.com/watch?v=oDDCoa71-yY>, consulté le 30/01/23.

n'offrant que du 8-bit (*chiptune*<sup>248</sup>), l'image et le son détournés restent en totale cohérence avec l'univers du jeu.

## 1.4 Visualisation et sonification des données

### 1.4.1 Traductions et transformations

Avec l'arrivée de nouvelles techniques liées au son ou à l'image, se développent quasi systématiquement des principes de traduction entre les deux composantes, par les intermédiaires d'un signal ou d'un codage. C'est ce que nous avons constaté avec le son optique dans les années 1930. Le signal d'amplitude étant inscrit graphiquement sur la pellicule, on pouvait en retour travailler ce motif graphique pour générer des notes de musique synthétique.

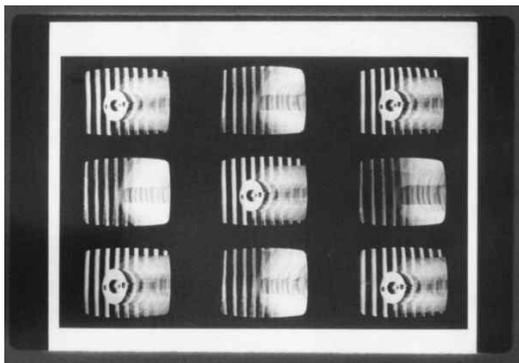


Fig. 54 : Steina et Woody Vasulka, *The Matrix*, 1970-1972 © Fondation Langlois

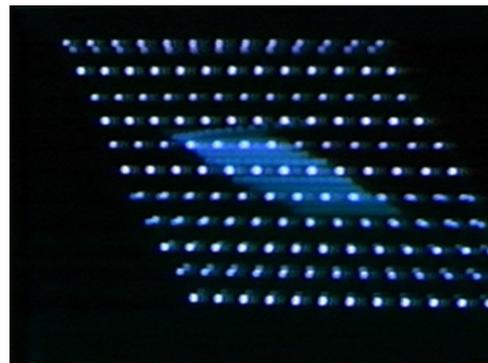


Fig. 55 : Steina et Woody Vasulka, *Soundsize*, 1974 © MoMA

Woody et Steina Vasulka sont certainement pour nous les artistes les plus représentatifs de ces explorations entre sonore et image dans les années 1970, car ils vont, inspirés par le travail de l'artiste Stephen Beck — artiste et inventeur du synthétiseur vidéo appelé Direct Video Synthetizer<sup>249</sup> — très vite utiliser les premiers synthétiseurs de son de l'époque, puis directement les oscillateurs (c'est-à-dire les générateurs d'ondes constitutifs de ces synthétiseurs) afin de fabriquer sur des moniteurs placés en matrice des images directement construites à partir de ces signaux.

<sup>248</sup> Le *chiptune* désigne une musique synthétique de faible définition, produite par une puce audio, souvent liée aux premiers jeux vidéo.

<sup>249</sup> Voir chapitre précédent, section 1.3.1.

Ils injectent différentes fréquences dans ces moniteurs afin d'en explorer les modes d'interférences. Au final, ils génèrent donc à la fois le son et l'image avec les mêmes signaux, ce qui va apporter une unité audiovisuelle jusque là très rare.

Nous nous rendions compte que la forme de base était la même pour l'audio et la vidéo : les ondes. Le wave-form generator était pratiquement le même pour l'audio et la vidéo. Nous étions donc en mesure d'envoyer ou de générer des images par le biais de signaux audio. Les différents arts avaient déjà depuis longtemps essayé d'unifier son et image, avec nos instruments nous étions enfin en mesure d'atteindre une véritable harmonie<sup>250</sup>.

Parmi les nombreuses œuvres qui utilisent ce procédé, citons *The Matrix* (1970-1972)<sup>251</sup>, une matrice de moniteurs au sein desquels s'exercent des relations entre les motifs graphiques abstraits et le son (comme le décrit Yvonne Spielmann, ils « génèrent le contenu audio de façon visuelle, produisant simultanément des images abstraites et sonores<sup>252</sup> ») et *Soundsizes* (1974)<sup>253</sup>, qui utilise un seul principe, minimaliste, celui de la transformation d'un plan constitué de points.

Les Vasulka vont également intégrer progressivement le numérique dans leur travail. À la fin des années 1970, Woody Vasulka, avec l'aide d'un physicien et d'un informaticien, va créer une machine nommée le *Digital Image Articulator*, « un instrument hybride conçu pour le traitement du signal vidéo qui conjugue des fonctions analogiques avec des composants de programmation numériques<sup>254</sup> ». Il permet « de décomposer l'image vidéo pixel par pixel et d'en articuler à nouveau la forme numérisée dans un environnement régi par les lois mathématiques<sup>255</sup> ».

---

250 Chris Dercon, « Entretien avec les Vasulka. La passion de la recherche », Bruxelles, in *Videodoc* n° 69, avril 1984, p. 17.

251 Une captation est disponible en ligne sur le site de la Fondation Langlois, <https://www.fondation-langlois.org/html/f/page.php?NumPage=495>, consulté le 30/01/23.

252 Yvonne Spielmann, « Steina et Woody Vasulka, *Matrix*, 1970-1972 », Fondation Langlois, 2004, <https://www.fondation-langlois.org/html/f/page.php?NumPage=495>, consulté le 30/01/23.

253 Une captation est disponible en ligne sur le site de la Fondation Langlois, <https://www.fondation-langlois.org/html/f/media.php?NumObjet=12620&NumPage=465>, consulté le 30/01/23.

254 Fondation Langlois, « Digital Image Articulator », <https://www.fondation-langlois.org/html/f/page.php?NumPage=457>, consulté le 08/07/22.

255 *Ibid.*

## 1.4.2 Face aux données : l'indifférence de l'ordinateur

Selon Bruno Bachimont, l'ordinateur fait preuve d'indifférence face aux données<sup>256</sup> :

[...] le numérique, en sa phase binaire, est un abstrait qui peut ignorer la matière dont il est réalisé. Il est indifférent à la matière physique dans laquelle il est implémenté. Il y a donc répétabilité à l'identique, indépendamment de la réalité des supports physiques<sup>257</sup>.

Cette indifférence est également décrite par le cinéaste et essayiste Malcolm Le Grice décrit dans un article :

À la capacité de synthétiser sans analyse s'ajoute celle de traduire les données. Étant donné que les données stockées dans la mémoire de l'ordinateur sont numériques, l'ordinateur est indifférent à la façon dont ces chiffres sont utilisés par un traitement de sortie. Les chiffres qui représentent les valeurs des couleurs des pixels dans une image de sortie peuvent prendre aussi facilement une forme qui rend des sons qu'une forme qui rend des images. Les données informatisées sont aussi aisément traduisibles que transformées, sans cependant que cela soit nécessairement significatif<sup>258</sup>.

Pour Bachimont le numérique — au sens de son code binaire — n'est abordable qu'à travers les transformations qu'on lui applique<sup>259</sup>, c'est-à-dire « des implémentations qui le réalisent et des interfaces qui le “phénoménologisent” (rendent perceptible et interprétable)<sup>260</sup> ». Les notions de transformation et de traduction sont donc inhérentes au numérique.

Dans ce que Lev Manovich a nommé l'esthétique des données, nous pourrions distinguer deux régimes de données<sup>261</sup>. Dans les systèmes numériques en effet, ce que l'on appelle la base de données est en général textuelle (base d'indexation), les médias

256 Nous l'avons évoqué brièvement dans l'introduction.

257 Bruno Bachimont, « Formes, concepts, matières : quels place et rôle pour le numérique et la technique », in *Technique et design graphique*, op. cit., p. 217.

258 Malcolm Le Grice, « L'abstraction chromatique : peinture, film, vidéo, image numérique », 1999, in Yann Beauvais (dir.), Malcolm Le Grice. *Le temps des images*, op. cit., p. 148-149.

259 Bruno Bachimont, op. cit., p. 225.

260 *Ibid.*, p. 227.

261 Les deux régimes peuvent s'entrecroiser. La géolocalisation des images par exemple, mixe un enregistrement du réel et un ensemble de données capturées (les coordonnées GPS ou l'heure d'enregistrement).

étant stockés en local (c'est-à-dire seulement accessibles à l'opérateur de l'interface qui en possède les droits d'accès). Elle contient également les métadonnées nécessaires à l'indexation du média enregistré (image, son, vidéo...). C'est le régime que nous étudierons dans la deuxième partie. Le média est stocké numériquement, compressé ou non, et attend son traitement. Cela est typique dans les logiciels de montage, dans lesquels on trouve une bibliothèque de médias stockés.

Le second régime de données informatiques serait celui non plus des données enregistrées (saisies), mais de celles capturées par des méthodes de mesure (la température par exemple). Récupérées sous une forme numérique brute, elles nécessitent, pour les matérialiser, des processus particuliers. Dans le cas de l'image et du son, les processus de « visualisation » ou de « sonification ».

Lev Manovich, lorsqu'il définit la visualisation des données, pose celle-ci comme un cas particulier du *mapping*<sup>262</sup>. Ce terme est communément utilisé pour définir la traduction de données en un autre type de données. Le terme reste toutefois ambigu<sup>263</sup>. Le théoricien des média Wolfgang Ernst le souligne bien en rappelant que l'étymologie du terme fait référence à la cartographie<sup>264</sup>. Une carte est une représentation qui permet la visualisation de et s'appuie sur des données relevées. L'autre sens du mot, lié aux mathématiques, est défini comme la traduction d'un jeu de données en un autre. En ce sens, le mot français dérivé de l'anglais, serait celui de « mappage » (une mise en correspondance de paramètres). Au sens étendu, il permet de circonscrire les frontières d'un processus informatique permettant de passer d'une représentation à une autre. La partition pourrait constituer un bon exemple, puisqu'elle est en quelque sorte un ensemble de données. Une partition imprimée déchiffrée et jouée par un interprète n'est pas un mappage informatique, car il y a une part d'interprétation. En revanche la partition traduite en données informatiques (au format MIDI par exemple), puis rejouée informatiquement est un exemple de mappage, la conversion des données

---

262 Lev Manovich, « Data Visualization as New Abstraction and as Anti-Sublime », 2002, <http://manovich.net/index.php/projects/data-visualisation-as-new-abstraction-and-anti-sublime>, consulté le 30/01/23.

263 Voir aussi le terme mapping vidéo.

264 Wolfgang Ernst, « Beyond the Archive: Bit Mapping », Media Art Net, 2004, [http://www.medienkunstnetz.de/themes/mapping\\_and\\_text/beyond-the-archive/1/](http://www.medienkunstnetz.de/themes/mapping_and_text/beyond-the-archive/1/), consulté le 30/01/23.

musicales en son musical. Ces données, si elles sont traitées de manière brute, pourraient automatiquement être traduites différemment, de manière générative.

Traduire un ensemble de données en un autre ensemble de données nécessite la mise en correspondance de paramètres concernant le premier ensemble, avec d'autres concernant le second ensemble. L'informatique permet la traduction des évolutions d'un nombre dans le temps en une entité graphique ; par exemple l'évolution du temps de lecture d'un son dans le lecteur d'une interface est généralement représentée par une barre de couleur dont l'origine se situe à gauche, et dont la largeur évolue en fonction du temps écoulé. Le temps total est lui défini par une barre fixe. Il s'agit d'une représentation normalisée, mais le pourcentage du temps écoulé pourrait tout à fait se manifester sous la forme d'un fond coloré passant progressivement du rouge au bleu, ou d'une accumulation progressive de motifs graphiques sur la page. Dans ces processus, c'est l'artiste qui va définir ses propres paramétrages de traduction.

### 1.4.3 Ryoji Ikeda, *test pattern*. La perception des données par l'hypersynchronisation audiographique

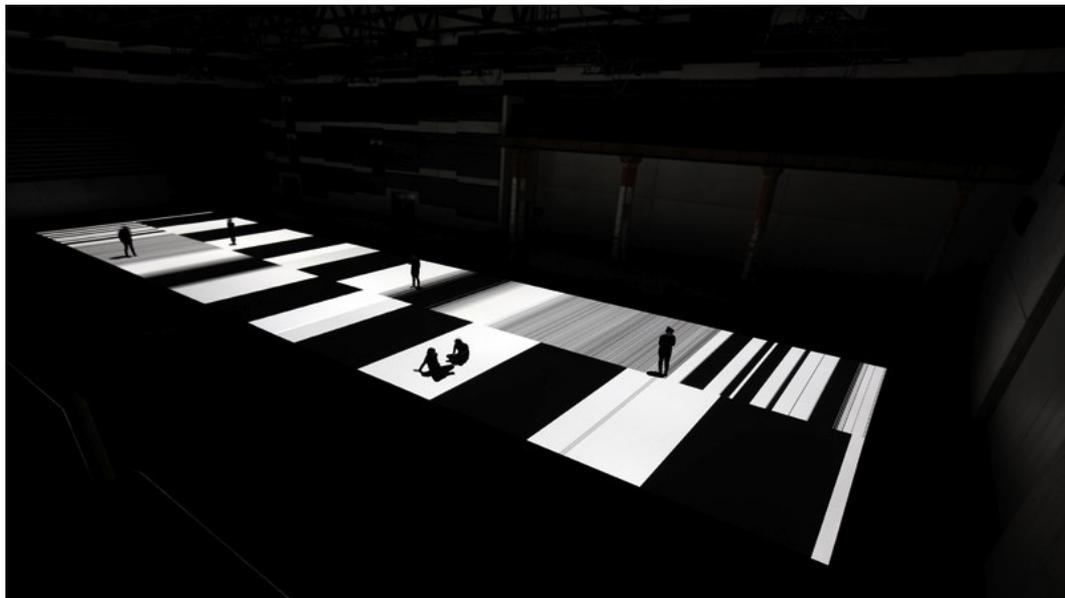


Fig. 56 : Ryoji Ikeda, *test pattern* [n° 5], installation audiovisuelle (2013)  
© Ryoji Ikeda / ryojiikeda.com

Au-delà des mappages formatés, les artistes vont chercher à se réappropriier ces processus et à créer leurs propres correspondances. L'artiste Ryoji Ikeda<sup>265</sup>, dont nous allons parler dans cette section, travaille cette indifférence de l'ordinateur dans des mappages audiovisuels radicaux. À ce titre, sa pièce audiovisuelle la plus emblématique est *test pattern*<sup>266</sup>.

Celle-ci existe sous plusieurs formes : disque CD, sculpture, performance, ou installation audiovisuelle. Cette dernière forme — dans ses versions monumentales récentes — consiste en une projection au sol de formes noires et blanches, assez proches de code-barres<sup>267</sup>. Sur une longue surface (en général plusieurs dizaines de mètres de longueur, selon les espaces) divisée en deux zones autonomes, ces ensembles d'aplats-lignes changent d'état à un rythme stroboscopique, tandis que des haut-parleurs disposés des deux côtés les plus longs de la surface mitraillent à une cadence similaire des sons électroniques minimalistes. Les sons sont à des tonalités très basses (infrabasses) ou très hautes (s'approchant des ultra-sons). Les visiteurs, à la fois spectateurs et auditeurs, circulent en marchant sur la projection (de très haute qualité).

### ***Un art adaptatif***

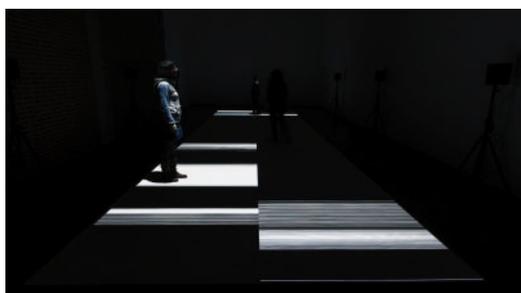


Fig. 57 : Ryoji Ikeda, *test pattern [n° 4]* (2013)  
© Ryoji Ikeda / ryojiikeda.com

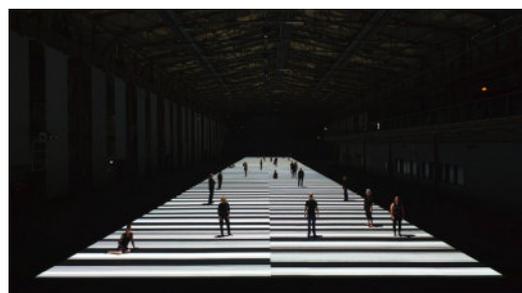


Fig. 58 : Ryoji Ikeda, *test pattern [100 m version]* (2013)  
© Ryoji Ikeda / ryojiikeda.com

Une des premières particularités de cette œuvre est d'être « déclinée » ou « explorée » sous plusieurs formes. Chacune de ces formes peut avoir plusieurs versions, qui dé-

265 Ryoji Ikeda est un artiste japonais qui a débuté dans les années 90 en tant que DJ. C'est notamment au sein du collectif multimédia Dumb Type qu'il commence à explorer les arts sonores, pour ensuite travailler sur ses propres créations sonores et ses performances qui sont à présent unanimement reconnues dans le monde de l'art. Ikeda est réputé pour ne pas donner d'interview ; il pense en effet que son œuvre doit parler d'elle-même.

266 Nous l'écrivons sans majuscules, comme le fait l'artiste.

267 Une captation d'une version de 100 m est disponible en ligne sur le Vimeo de Ryoji Ikeda, <https://vimeo.com/120454782>, consulté le 30/01/23.

pendent du contexte de présentation, ou bien de commandes d'institutions rendant l'évolution de l'œuvre possible. Si nous nous en tenons au site officiel de l'artiste, il en existe actuellement huit versions principales : 13 versions de l'installation audiovisuelle, une version performance, une version sur pellicule, et enfin un CD de 16 pistes. Dans une analyse du travail de Ryoji Ikeda en 2012, le directeur artistique et commissaire en chef du YCAM<sup>268</sup> Kazuneo Abe imaginait que « [...] si potentiellement Ikeda avait à sa disposition un espace de 1000 m de long, il étendrait sûrement l'installation afin qu'elle en couvre toute la surface<sup>269</sup> ». En 2013, Ikeda proposait déjà une version de *test pattern* sur 100 m de long, projetée au sol et l'on comprend bien que cette œuvre est encore amenée à évoluer. En 2014, *test pattern* a même été projeté sur les écrans de Times Square<sup>270</sup>.

Sur ces aspects liés au principe de « version », Lev Manovich nous fournit plusieurs concepts tirés de son analyse des nouveaux médias<sup>271</sup>, tels que la modularité, l'automatisation et la variabilité d'une œuvre. La variabilité est justement directement liée à l'idée de version, l'automatisation définit quant à elle l'autonomie et la capacité de l'œuvre à être jouée à l'infini ou en boucle. Le cinéaste et théoricien Yann Beauvais nous parle très bien de la modularité.

Par exemple, la modularité montre bien comment les nouveaux médias possèdent la même structure à différents niveaux. Un objet médiatique est constitué de parties indépendantes qui sont accessibles, modifiables et combinables, sans pour autant perdre leur identité ; et ceci à différentes échelles<sup>272</sup>.

Nous sommes ainsi en présence d'une œuvre contextuelle, qui peut être « modulée ». Les progrès technologiques rapides et les évolutions de la projection, des écrans, amènent les artistes à proposer des travaux qui pourront « survivre » en s'adaptant à leur environnement. C'est un enjeu important et passionnant pour les œuvres basées sur la technologie. Dans le design web, on constate d'ailleurs une tendance actuelle au

268 YCAM Yamaguchi Center for Arts and Media, <http://www.ycam.jp/en/>, consulté le 30/01/23.

269 Kazuneo Abe, « As “/” », in Ryoji Ikeda, *Datamatics*, Milan, Charta, 2012, p. 107.

270 Cette projection avait eu lieu dans le cadre d'un événement artistique, Midnight Moment, qui donne à des artistes l'accès aux écrans publicitaires pendant 3 minutes, juste avant minuit, chaque jour pendant un mois. Pour écouter le son de *test pattern*, les spectateurs pouvaient emprunter un casque, pour avoir le son synchronisé.

271 Yann Beauvais, « Code source ouvert », in Lev Manovich, *Le langage des nouveaux médias*, op. cit., p. 11.

272 Yann Beauvais, « Code Source ouvert », in Lev Manovich, *Le langage des nouveaux médias*, op. cit., p. 9.

« responsive design<sup>273</sup> » (design adaptatif), qui consiste à créer des design de pages capables de se recomposer, de se transformer selon la taille de l'écran utilisé. Cela permet à un même design d'être vu correctement aussi bien sur un mobile que sur un écran de grande taille. À la différence du peintre qui choisissait son cadre avant de composer son œuvre, l'artiste compose une œuvre susceptible d'être projetée ou vue sur différentes surfaces.

### ***Seuils de perception audio-visuelle***



Fig. 59 : Ryoji Ikeda, *test pattern [n° 1]*, installation (2008)  
© Ryoji Ikeda / ryojiikeda.com

Dans *test pattern n° 1*, la première version de l'installation, les codes-barres sont générés à partir du son, par un processus de conversion. Pour les versions ultérieures, Ikeda semble utiliser un logiciel sur mesure, programmé par Tomonaga Tokuyama. Celui-ci permet à l'artiste de charger n'importe quel type de fichiers (texte sons, photos ou films), pour les convertir en données binaires (des 0 et des 1). À partir de ces chiffres, il va générer des suites de lignes blanches (pour le 0) et noires (pour le 1) s'apparentant à des code-barres à l'écran et des sons. L'image d'une part, et le son, d'autre part, seraient ainsi générés par un même code de base. Étonnamment, malgré le peu de moments de calme dans la diffusion, une sensation de plénitude s'installe au bout d'un

<sup>273</sup> Le terme « responsive » avait notamment été utilisé en 1965 dans le titre de l'exposition du MoMA de New York qui a consacré l'Op'art : *The Responsive Eye*, MoMA, New York, du 25 février au 25 avril 1965. Commissaire : William C. Seitz.

certain temps, à la manière de la *Dream House* (1964) de La Monte Young et Marian Zazeela<sup>274</sup>.

Selon le texte de présentation de la pièce<sup>275</sup>, le projet est destiné à explorer la relation entre les limites de la machine et les seuils de la perception humaine. Le titre de l'œuvre est à cet égard tout à fait cohérent, car ce que l'on nomme « test pattern » ou « test card » est un signal de test utilisé pour les appareils qui nécessitent un calibrage particulier, comme les téléviseurs.

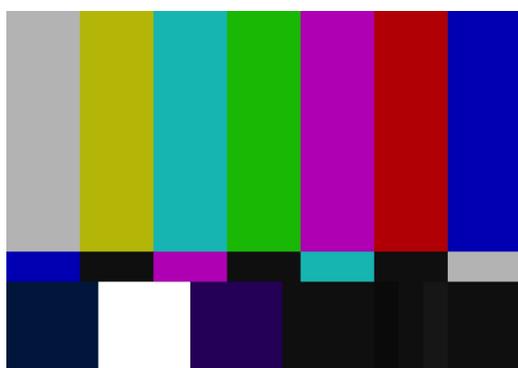


Fig. 60 : SMPTE color bars  
© Wikipedia

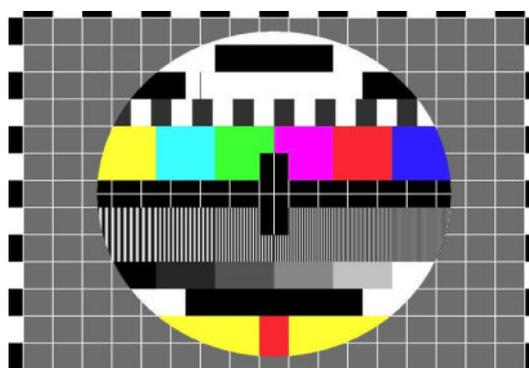


Fig. 61 : Philips PM5544 (SECAM variation)  
© Wikipedia

L'ouvrage *Testing Hearing*<sup>276</sup>, d'Alexandra Hui, Mara Mills et Viktoria Tkaczyk, est une historiographie des tests d'écoute. Il débute, tout à fait pertinemment, par la description de *test pattern*. Les autrices signalent que sur la version CD — qui préexiste à l'installation audiovisuelle —, une mention précise qu'un fort volume peut endommager l'audition de l'utilisateur. C'est bien ce qu'évoque Ikeda, par le titre de l'œuvre qui fait référence aux signaux audiovisuels de test pour les télévisions. Les autrices font également un pont entre ce travail et l'œuvre pionnière de John Cage, *Imaginary Landscape n° 1* (1939), qui utilise des disques de test pour les studios. Ces disques contiennent des signaux sonores primaires (des sons sinusoïdaux par exemple). Là encore, la référence

274 Dans un registre similaire, la *Dream House* (1964 —) est un espace immersif de lumière et de son imaginé par le compositeur expérimental La Monte Young et la plasticienne Marian Zazeela. Les deux artistes avaient remarqué que les corps, exposés constamment à des sons (des drones sonores) et à la lumière magenta, finissaient en quelque sorte par s'accorder à son environnement.

275 Page de présentation de *test pattern*, <http://www.ryojiikeda.com/project/testpattern/>, consulté le 30/01/23.

276 Viktoria Tkaczyk, Mara Mills, and Alexandra Hui (dir.), *Testing Hearing*, Oxford, Oxford University Press, 2020.

est pertinente, puisque Ryoji Ikeda explique avoir trouvé ses axes de travail à partir du moment où il s'est posé la question de ce qui était fondamental dans le son<sup>277</sup> : pour lui, l'onde sinusoïdale (un son dit pur) et le bruit blanc (son qui contient toutes les fréquences du spectre). À ce titre nous pouvons également établir une analogie forte avec le film structurel *Arnulf Rainer* (1960), réalisé par le cinéaste Peter Kubelka. Celui-ci joue avec les fondamentaux du cinéma : la lumière et l'obscurité, le silence et le bruit (blanc). Par la radicalité d'emploi de ces composantes, il rééquilibre le rapport entre image et son, qui deviennent égaux. Son utilisation du *flicker* (que l'on traduirait par clignotement ou scintillement) pose aussi le cinéma non pas comme un art du mouvement, mais comme une projection rapide d'impulsions lumineuses :

Le cinéma n'est pas mouvement. C'est la première chose. Le cinéma n'est pas mouvement. le cinéma est une projection d'images fixes — ce qui veut dire des images qui ne bougent pas — à un rythme très rapide. Et vous pouvez, bien sûr, provoquer l'illusion d'un mouvement, mais c'est un cas particulier, et le cinéma a été inventé à l'origine pour ce cas particulier<sup>278</sup>.

Nous pouvons ainsi parler, avec *test pattern*, d'une « hypersynchronisation » audio-graphique (la cadence dépasse les 100 images par seconde). Il y est difficile d'appréhender l'image ou le son indépendamment. C'est une sorte d'appairage, de couplage extrême entre l'image et le son, tous deux diffusés sous une forme brute, radicale et élémentaire.

---

277 « Patterns of the Unknown », entrevue avec Marcella Lista, in Ryoji Ikeda, *Continuum*, Paris, Xavier Barral/Centre Pompidou, 2018, p. 156.

278 Christian Lebrat, *Peter Kubelka*, Paris, Paris Expérimental, 1992, p. 61.

### 1.4.4 La sonification. Des sciences à l'art



Fig. 62 : Différents modèles de stéthoscopes de Laënnec  
© Wellcome Collection

Vers 1816, le physicien français René Laënnec imagine une méthode d'auscultation (en latin *auscultare* : écouter<sup>279</sup>) basée sur un outil cylindrique en bois<sup>280</sup> : le premier stéthoscope. Dans son ouvrage *De l'auscultation médiate*<sup>281</sup> (1819), il catégorise les différentes familles de bruits du thorax (pour les battements de cœur : ses irrégularités, les intermittences de rythme, les palpitations...). Le son devient un outil de diagnostic pour la médecine. Avec l'invention du téléphone et du phonographe, ce qui n'est alors que de l'amplification (le stéthoscope) va être médiatisée par des appareils de conversion directe et ouvrir une première époque de la sonification<sup>282</sup>. Au cours du XX<sup>e</sup> siècle, de l'émergence d'appareils médicaux sonores jusqu'aux instruments de détection comme le compteur Geiger, le détecteur de métaux ou le sonar — qui génèrent du son pour signaler une présence —, les sciences trouvent des utilités inédites au sonore.

279 Florian Grond et Theresa Schubert-Minski, « Sonification: Scientific method and Artistic Practices », in *See This Sound*, *op. cit.*, p. 286.

280 Les auscultations étaient auparavant réalisées directement à l'aide de tapotements directs sur le corps, par l'analyse des résonances naturelles.

281 René Laënnec, *De l'auscultation médiate, ou Traité du diagnostic des maladies des poumons et du cœur, fondé principalement sur ce nouveau moyen d'exploration*, Paris, Brosson, 1819.

282 Florian Grond et Theresa Schubert-Minski, *op. cit.*, p. 286.

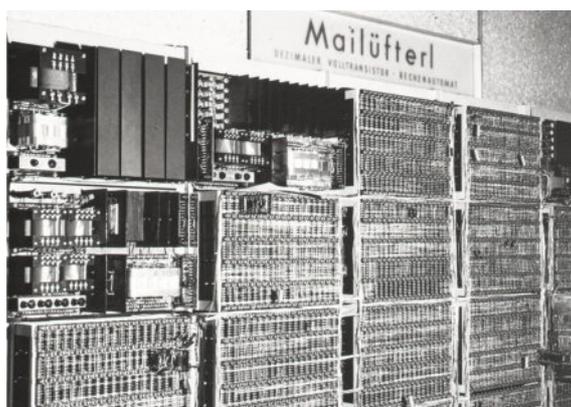


Fig. 63 : Le Mailüfterl de Heinz Zemanek  
© TU Wien Archive/Heinz Zemanek Estate

Avec l'informatique, les possibilités s'étendent. Dans leur article consacré aux pratiques de sonification<sup>283</sup>, Florian Grond et Theresa Schubert-Minski relatent une expérience étonnante. Heinz Zemanek, pionnier de l'informatique, avait inventé, à la fin des années 1960, un ordinateur qu'il avait nommé Mailüfterl. La particularité de cette machine était qu'elle était couplée à un générateur de son à douze tonalités, ce qui lui permettait de générer des séquences sonores. Quand tout allait bien, la machine émettait une séquence mélodique. Si des erreurs de traitement informatique se produisaient, alors la série de sons n'était plus harmonieuse, et les ingénieurs savaient qu'il se passait quelque chose d'anormal. Comme les tâches de traitement étaient souvent longues, et qu'il n'était pas concevable de rester en permanence à écouter la machine, le Mailüfterl avait été connecté au téléphone. Les ingénieurs appelaient régulièrement, de n'importe où, pour connaître l'avancement du travail grâce au son.

Le terme « sonification » n'est utilisé qu'à partir de 1992, avec Gregory Kramer et l'International Community for Auditory Display (ICAD). La définition, souvent reprise, de Kramer, est la suivante :

La sonification est l'utilisation de signaux sonores non verbaux permettant de véhiculer de l'information. Plus précisément, la sonification est la transformation de relations de données afin que celles-ci deviennent perceptibles sous

---

283 *Ibid.*, p. 287.

forme sonore dans le but d'en faciliter la communication ou l'interprétation [Traduction de Lorella Abenavoli]<sup>284</sup>.

Les pratiques artistiques de sonification sont toutefois antérieures à l'utilisation du terme. Citons par exemple, en musique, le compositeur Charles Dodge qui a réalisé en 1970 une composition de musique informatique en se basant sur un tableau des indices kp (mesures de la force moyenne du champ magnétique terrestre). Dodge s'est basé sur les données d'une année, soit près de 3000 nombres. Grâce à un principe de mappage, il a lié proportionnellement la force en kp à la hauteur musicale (sur quatre octaves), puis il a créé un programme qui a synthétisé cette suite de sons, pour obtenir 8 minutes de musique.

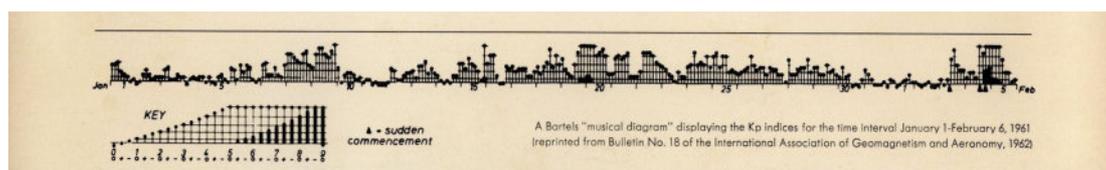


Fig. 64 : Tableau de données utilisé par Charles Dodge  
© UMKC Special Collections

Le scientifique, en général, recueille, analyse et visualise des données grâce à des appareils de captation, à l'échelle microscopique ou macroscopique. Le design d'information permet de chercher à faire varier les manières de représenter des phénomènes.

De plus en plus d'artistes collaborent avec les scientifiques dans l'idée d'utiliser, voire de détourner ce design de données en vue de fabriquer des traductions singulières (non modélisantes, selon les mots de Johanna Drucker<sup>285</sup>) et critiques. Certains opèrent dans un même système artistique des visualisations et des sonifications, créant de nouveaux rapports audiovisuels. Que ce soit à l'échelle des mouvements respiratoires

284 « Sonification is the use of nonspeech audio to convey information. More specifically, sonification is the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation » (Lorella Abenavoli, *Le son plastique : empreindre le flux et l'inouï. Sonification et audification dans l'art de l'installation*, sous la direction de Louise Poissant, Université du Québec à Montréal, 2017, p. 24).

285 C'est-à-dire des traductions qui ne cherchent pas à représenter des informations en suivant les modèles imposés par les logiciels et leurs interfaces (Johanna Drucker, *Visualisation. L'interprétation modélisante*, Paris, B42, 2020).

avec Jacopo Baboni Schilingi, de ceux de la terre avec Semiconductor, ou des flux éoliens avec Peter Beyls, l'expérience artistique des données s'enrichit avec l'audiovisualité.

### 1.4.5 Audiographie du corps humain

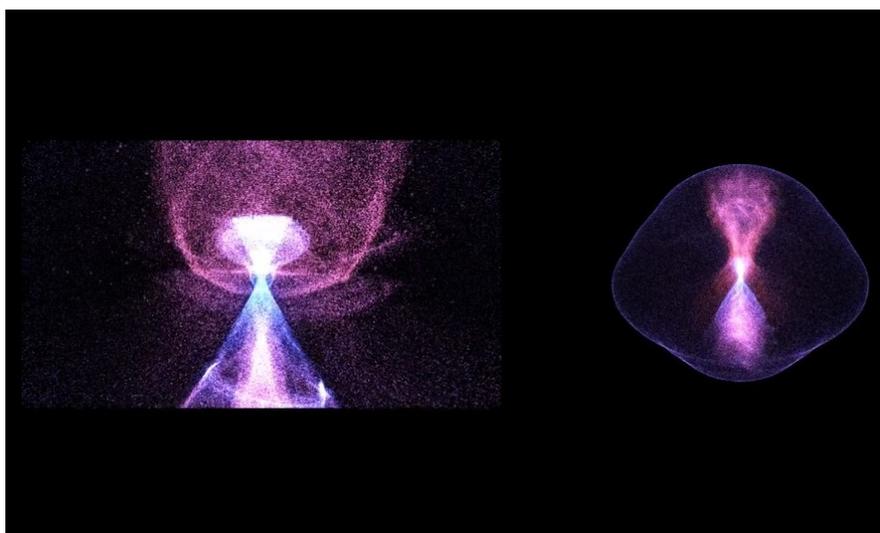


Fig. 65 : Jacopo Baboni Schilingi, *Argo* (2018)  
© Claude Mossessian / Grand Palais

Le stéthoscope permettait pour la première fois d'amplifier les bruits du thorax. Depuis, nos appareils mobiles enregistrent numériquement, les données de nos corps. Grâce à leur podomètre intégré, ils mesurent le nombre de pas que nous faisons dans la journée. Avec *Argo*<sup>286</sup>, installation générative de Jacopo Baboni Schilingi, il s'agit de respiration.

L'artiste porte en permanence un capteur de respiration conçu par un ingénieur spécialisé dans la mesure de la respiration. Les dilatations de la cage thoracique et les positions du corps de l'artiste dans l'espace sont récupérées, puis visualisées et sonifiées en temps réel. *Argo* peut récupérer les données à n'importe quel endroit de la planète, les transforme en données musicales et données visuelles<sup>287</sup>.

<sup>286</sup> *Argo* était le chien d'Ulysse, qui l'a attendu pendant son odyssee. Baboni Schilingi réinterprète cette métaphore de l'attente en nommant son système *Argo*, qui, depuis le serveur qui l'héberge, récupère à distance les données de l'artiste constamment en déplacement dans le monde.

<sup>287</sup> Une captation est disponible en ligne sur le Vimeo de la galerie Charlot, <https://vimeo.com/345013329>, consulté le 30/01/23.

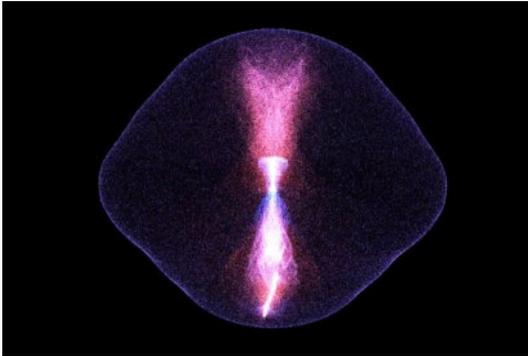


Fig. 66 : Jacopo Baboni Schilingi, *Argo* (2018)  
Une inspiration, en rouge  
© Claude Mossessian / Grand Palais

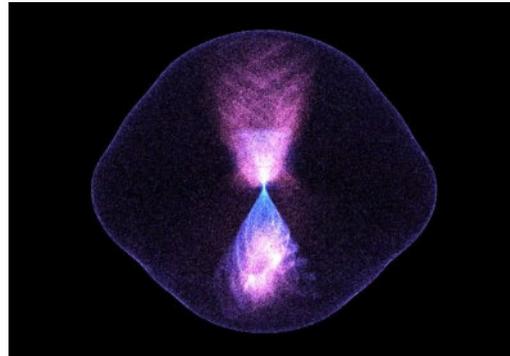


Fig. 67 : Jacopo Baboni Schilingi, *Argo* (2018)  
Une expiration, en bleu  
© Claude Mossessian / Grand Palais

La représentation graphique animée du thorax en temps réel, même si elle est une abstraction, permet de deviner de véritables événements corporels, sous forme par exemple de flux granulaires : les inspirations de l'artiste sont en rouge, ses expirations sont en bleu.

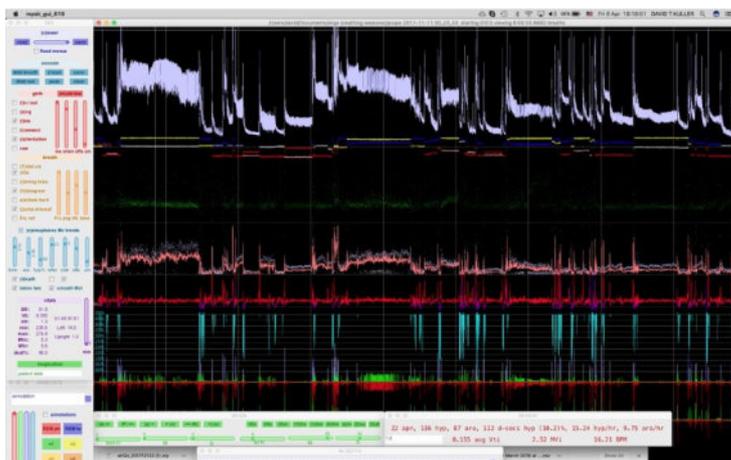


Fig. 68 : Jacopo Baboni Schilingi, *Argo* (2018)  
Visualisation traditionnelle des données respiratoires récupérées  
© Claude Mossessian / Grand Palais

Le schéma des données récupérées en permanence, est susceptible, pour un spécialiste de montrer quels sont les phases de sommeil ou les déplacements importants.

### 1.4.6 Géologie vidéo-sonique : *Earthworks*, de Semiconductor

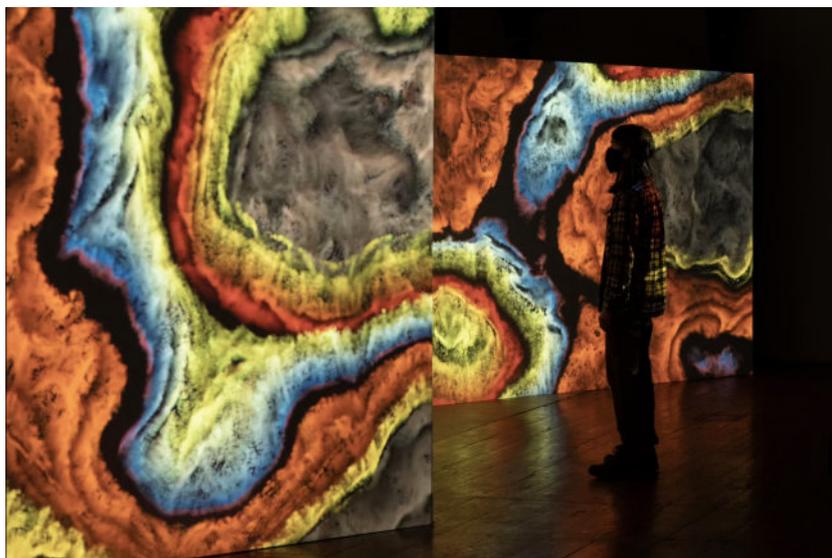


Fig. 69 : Semiconductor, *Earthworks* (2016)  
© Semiconductor

Les mouvements naturels du sol de notre planète se situent dans des temps géologiques extrêmement longs, ce *Deep Time* qui « permet de penser l'histoire de la Terre indépendamment de tous repères historiques (humains) »<sup>288</sup>. Imperceptibles à l'œil nu, les phénomènes géologiques naturels contrastent avec ceux, multiples et artificiellement causés par l'homme. Dans l'installation *Earthworks* (2016)<sup>289</sup>, le duo britannique Semiconductor<sup>290</sup> cherche à montrer et à faire entendre ces mouvements.

Dans l'espace, cinq écrans juxtaposés en zigzag composent un flux horizontal continu d'images. Sans aucune autre information à l'écran, des strates colorées de matière granulée ondulent, au rythme de sons de craquement et de crissements inquiétants. Différents tableaux animés se succèdent ainsi. Des moments calmes s'enchaînent avec ce qui pourrait s'apparenter à de brefs éboulements, dans des vagues de mouvements qui semblent « vivants ».

288 Pierre Savaton, « Le temps fondateur de la géologie », in *Romantisme* 2016/4 (n° 174), p. 39.

289 Une version est disponible en ligne sur le site de Semiconductor, <https://semiconductorfilms.com/art/earthworks/>, consulté le 30/01/23.

290 Semiconductor travaille depuis les années 2000 avec des données scientifiques, qu'ils détournent et modélisent numériquement, par le son et par l'image.

Au premier abord, et comme nous le laisse supposer le titre de la pièce, *Earthworks*, ces flots audio-visuels abstraits semblent évoquer les mouvements du sol. Ils sont pourtant beaucoup plus qu'une évocation.

Pour ce travail, Semiconductor a collaboré avec un institut scientifique pour récupérer des jeux de données relatives à différentes natures de mouvements géologiques : l'activité sismique glaciaire, l'activité sismique terrienne (les tremblements de terre), l'activité sismique organique (les magmas volcaniques). L'IRIS (Incorporated Research Institutions for Seismology) leur a ainsi donné un accès à un ensemble d'archives de sismogrammes, qui reflètent des mouvements du sol. Un sismogramme est un enregistrement graphique de données sismiques dans un temps donné. Comme le pointe Ruth Jarman, l'une des membres de Semiconductor, le signal graphique qui en résulte est une forme d'onde, tout à fait semblable dans sa morphologie au signal d'amplitude d'un son<sup>291</sup>.

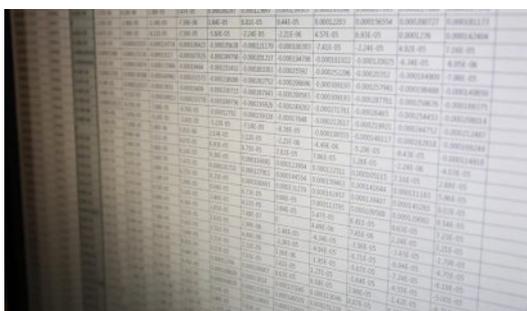


Fig. 70 : Semicondutor, *Earthworks* (2016)  
Données brutes  
© Semiconductor



Fig. 71 : Semicondutor, *Earthworks* (2016)  
Représentation graphique des données  
© Semiconductor

Les artistes ont sélectionné, avec l'institut, des jeux significatifs (par dates et lieux), et ont directement converti ces données en son. Jarman insiste sur le fait que ce ne sont pas des données qui suggèrent ou évoquent, mais bel et bien les données brutes lisibles directement sous forme de bruits<sup>292</sup>. Les fichiers d'onde sont tout simplement

291 Le signal d'amplitude d'un son est la représentation graphique de l'intensité de son amplitude dans le temps. Un silence est donc représenté par un vide. Il s'agit de la représentation du son la plus utilisée, qui est celle affichée par défaut lorsque l'on ouvre un fichier son dans un logiciel de traitement sonore.

292 *Earthworks* by Semiconductor (2020), film documentaire de Tom Thistlethwaite pour la Fabrica Brigh-ton, <https://www.youtube.com/watch?v=xkg38zms44Q&t=255s>, consulté le 30/01/23.

ouverts (quelques ajustements spécifiques sont nécessaires) avec un logiciel de traitement sonore différent que celui utilisé par les scientifiques concernés. Il s'agit, comme nous l'avons déjà évoqué, d'utiliser l'indifférence de l'ordinateur pour lire des données non sonores dans un logiciel sonore : une sonification directe, ce que des théoriciens de la sonification appellent l'audification<sup>293</sup>.

Ces données se font sur des durées importantes. Les artistes voulaient à la fois mettre en scène ces mouvements longs, mais surtout les mettre en regard avec ceux qui résultent de l'activité humaine, comme les excavations multiples menées pour l'extraction de nos matières premières. Ils ont donc, avec l'aide de scientifiques de l'Université de Barcelone, enregistré des ondes sismiques similaires à celles, naturelles, des archives existantes, mais dans une carrière d'extraction active, près de la ville espagnole de La Planta. On peut ainsi, comme le résume Joe Gerhard — l'autre membre du duo —, « littéralement entendre l'humain modifier le sol de la planète<sup>294</sup> ».

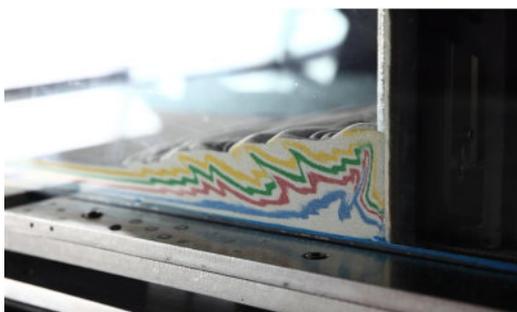


Fig. 72 : Semiconductor, *Earthworks* (2016)  
Exemple de géomodèle en couches  
© Semiconductor



Fig. 73 : Semiconductor, *Earthworks* (2016)  
Enregistrement d'un sismogramme  
© Semiconductor

Ces quatre ensembles de sons (les trois relatifs aux données sismiques, et celui de l'activité humaine) constituent les bandes sonores respectives des quatre animations qui vont se succéder à l'écran. Là encore, plus qu'une illustration sonore plaquée, il s'agit pour les artistes de faire vivre visuellement ces données sonores. Ils ont donc préparé, pour chacun des quatre ensembles sonores, une animation CGI (*Computer-Generated Imagery*), c'est-à-dire une animation de synthèse. Le travail visuel se fonde sur une

293 Thomas Hermann, Andy Hunt, John G. Neuhoff (dir.), *The Sonification Handbook*, Berlin, Logos Publishing House, 2011, p. 17.

294 *Ibid.*

technique utilisée par les scientifiques depuis le milieu des années 1850, la modélisation analogique. Il s'agit d'empiler des couches de sable coloré dans un réceptacle, puis de soumettre cet ensemble à des forces physiques (poussées, compressions...) exercées humainement ou mécaniquement. Les animations visuelles qui en résultent sont des abstractions visuelles, en haute définition, composées de strates colorées qui ondulent.

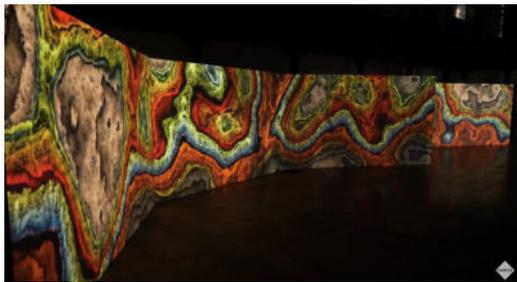


Fig. 74 : Semiconductor, *Earthworks* (2016)  
Activité sismique organique  
(les magmas volcaniques)  
© Semiconductor

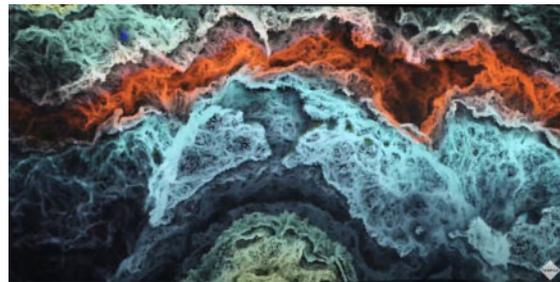


Fig. 75 : Semiconductor, *Earthworks* (2016)  
Activité sismique glaciaire  
© Semiconductor

Au-delà de l'aspect visuel, il faut bien comprendre que l'animation n'est pas conçue de manière linéaire. Elle est préparée pour réagir à des données extérieures, plus précisément ici les données sonores. Elle est en quelque sorte branchée sur le son. S'il n'y a pas de bruit, elle reste immobile. Elle est une visualisation directe du son. Il y a donc un paramétrage à opérer, plus précisément un mappage programmé informatiquement, qui indique à l'animation de se mettre en mouvement selon des paramètres sonores détectés : par exemple, le volume ou la fréquence.

Il y a donc un couplage fort opéré entre le son et l'image. Si l'on entend des tremblements sonores, on voit les strates colorées tremblées. Il y a une corrélation physique, de mouvement, entre l'image et le son, ce qui crée une vraie physicalité (qui manque souvent aux animations synthétiques). Le fait que le son soit une traduction directe, et que la partie visuelle réactive se fonde sur un modèle géologique proche de notre imaginaire du sol provoque chez le visiteur la sensation d'une présence audiovisuelle.

Plusieurs artistes ont utilisé la technique d'audification des ondes sismiques. La plasticienne Lorella Abenavoli réalise, dans les années 1990 et 2000, l'œuvre *Le Souffle de la Terre*, en collaboration avec des géophysiciens et des ingénieurs. Elle développe le logiciel SdT (Son de la Terre), qui permet de transformer les ondes infragraves de la Terre en son<sup>295</sup> sur un principe d'audification. Citons également Florian Dombois avec *Circum Pacific 5.1* (2001), une installation sonore à 5 canaux qui permet d'écouter les ondes sismiques à différents endroits du globe.

### 1.4.7 Peter Beyls. Des systèmes audiographiques auto-organisés

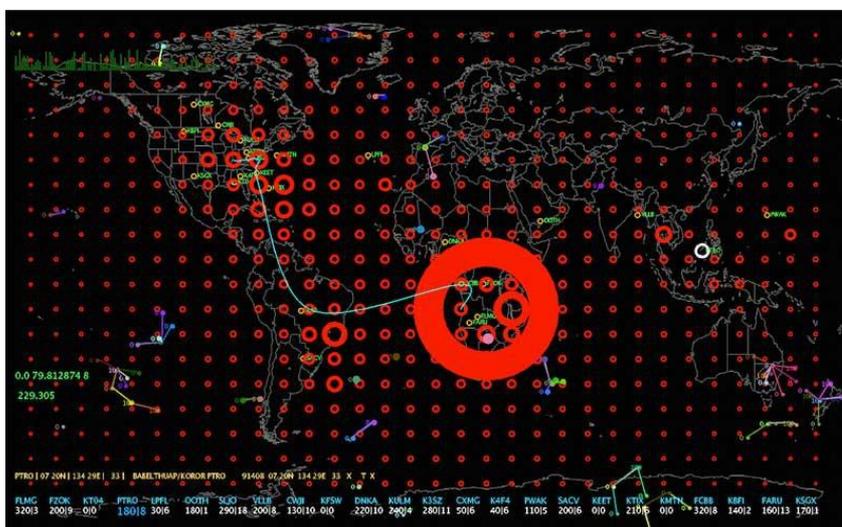


Fig. 76 : Peter Beyls, *WindChime* (2006-2017)  
Installation audiovisuelle en temps réel à partir de données en ligne  
© Galerie Charlot

L'utilisation de données peut se faire de manière figée, c'est-à-dire que les données sont traitées numériquement une fois pour toutes, puis mises en forme : une image, une représentation 3D, etc. L'accessibilité croissante d'ensemble de données brutes (des listes de nombres) de nature diverse (scientifiques, sociales, etc.) offerte par le numérique nourrit l'éventail de potentialités de matérialisation de celles-ci, d'un rendu direct et lisible à un détournement total.

295 « Lorella Abenavoli. Le Souffle de la terre », Fondation Langlois, <https://www.fondation-langlois.org/html/f/page.php?NumPage=268>, consulté le 30/01/23.

L'artiste Peter Beyls travaille sur ces relations paradoxales en les croisant dans des systèmes programmés. Il a d'abord été musicien et compositeur, dans les domaines de la musique sur bande et électronique. Il a dans ce contexte travaillé lui-même sur des systèmes de manipulation en temps réel de l'électronique, mais également sur le développement d'un synthétiseur<sup>296</sup>.

Ses premiers travaux visuels sont en majorité des œuvres abstraites, géométriques, générées par ordinateur et imprimées sur table traçante. La potentialité conceptuelle des programmes qu'il conçoit fait qu'il s'intéresse à la fin des années 80 aux systèmes cellulaires et aux agents artificiels.

Beyls conçoit les médias informatiques comme des partenaires actifs dans un processus créatif, une méthodologie qu'il appelle « navigation conceptuelle ». Le logiciel est écrit dans le but d'explorer des intentions ambiguës. Une fois qu'une idée est formalisée dans un programme, on peut évaluer son potentiel imaginaire grâce à la rétroaction qu'il fournit. Puisqu'un programme reflète les objectifs de l'artiste, la programmation est considérée comme une méthode d'introspection esthétique. Le logiciel est donc instrumental en tant que moyen fonctionnel et matérialiste permettant la manipulation active de constructions par ailleurs purement conceptuelles<sup>297</sup>.

L'une de ses pièces, *Windchime* (2006-2017)<sup>298</sup>, est une pièce audio-graphique générative connectée dans laquelle se recoupent bon nombre de ses préoccupations liées aux systèmes et à la dynamique des données.

Beyls s'intéresse dans ce travail à la dynamique des vents, celle des flux éoliens qui circulent et se croisent dans la stratosphère. Le vent n'est perceptible que par ses effets, le mouvement des feuilles d'un arbre par exemple, qui en matérialisent ses mouvements. Le son peut également en constituer un indicateur intéressant, Beyls prend

---

296 Galerie Charlot (exposition *Sound Vibes*), <https://www.galeriecharlot.com/fr/79/Peter-Beyls>, consulté le 30/01/23.

297 Galerie Charlot, texte de présentation de l'exposition *Sound Vibes*, 2019, <https://www.galeriecharlot.com/fr/79/Peter-Beyls>, consulté le 30/01/23.

298 Une documentation vidéo présentée par Peter Beyls est disponible sur le Vimeo de l'Université catholique de Porto, <https://vimeo.com/128248373>, consulté le 30/01/23.

comme référence le carillon à vent (*Windchime*), mobile musical suspendu dont les cylindres s'entrechoquent et composent une mélodie percussive.

*Windchime* est une installation audiovisuelle connectée. Ses composantes visuelles et sonores naissent de l'extraction et de la matérialisation en temps réel de données issues d'une base météorologique, parmi presque 8000 lieux disséminés sur le globe. Beyls considère le monde comme un réservoir de données, un *Found System* selon ses propres termes<sup>299</sup>. Cette expression fait référence au *Found Footage*, au sens esthétique, celui de réemploi ou emploi ; il désigne la réappropriation de matériaux plastiques déjà existants dans un autre contexte. Les systèmes présents dans l'Univers, qu'ils soient sociaux, biologiques, cellulaires ou galactiques<sup>300</sup>, ne sont pas directement observables. Par les données, ils peuvent être explorés à des fins esthétiques, par toutes les matérialisations possibles ; ici par le mouvement d'abstractions géolocalisées et de générativités sonores.

Le seul écran de l'installation nous donne à voir une mappemonde, fixe, sur laquelle une trame constituée de petits disques évolue, certains disques s'activant à un moment donné. Autour de ces disques, des particules s'animent, et visualisent les intensités et les directions des flux éoliens. Dans le même temps, des sons synthétiques sont joués, selon une logique imperceptible d'emblée.

Techniquement, le programme central, écrit en langage Java<sup>301</sup>, communique en XML<sup>302</sup> (un format d'échange de données très répandu) avec un serveur du National Center for Atmospheric Research (NCAR), structure dont la mission est d'explorer scientifiquement les couches atmosphériques. Le programme produit ensuite les données indirectes qui sont envoyées à un moteur de génération de son synthétique écrit en SuperCollider<sup>303</sup>, envoyé par OSC<sup>304</sup>, une interface de transmission. Les langages et

299 Peter Beyls, « WindChime: database driven mapping Of Complex dynamical systems », 2014, <http://2014.xcoax.org/pdf/xcoax2014-Beyls.pdf>, consulté le 30/01/23.

300 *Ibid.*

301 Langage de programmation orienté objet créé par James Gosling et Patrick Naughton en 1995.

302 Extensible Markup Language file, développé par les équipes de John Bosac et sorti en 1998.

303 Environnement et langage de programmation dédié à la synthèse audio en temps réel et à la composition algorithmique, créé en 1996 par James McCartney.

304 Open Sound Control, un protocole de transmission audio entre appareils compatibles.

protocoles utilisés ici sont assez standards, la complexité réside surtout dans la conceptualisation et l'écriture du programme.

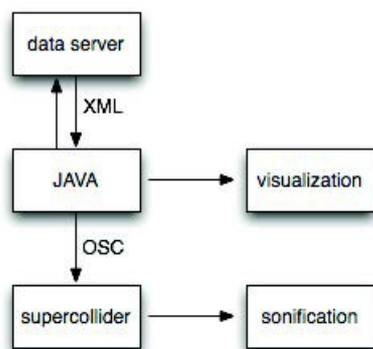


Fig. 77 : Peter Beyls, *WindChime* (2016-2017), Schéma de circulation des données © Peter Beyls, « WindChime: database driven mapping Of Complex dynamical systems », 2014

À intervalles réguliers, le programme analyse les données de 24 lieux (les stations), tirés aléatoirement parmi 7961, puis normalise celles-ci pour générer une animation faisant apparaître visuellement les fluctuations de données. Le son est produit par des générateurs de son sinusoïdal, qui peuvent produire, par leur combinaison et leur non-linéarité, des spectres audio complexes. Il y a 4 synthétiseurs virtuels, chacun étant lié à un type de particule différent. Il existe une dépendance aux « voisins ». Une particule isolée (sans voisins) ne produira pas de son. L'accumulation d'énergie produira plus de flux audio.

Dans la pièce de Beyls, deux systèmes génératifs interagissent. L'un, visuel, est directement généré à partir des listes de nombres et des localisations réelles ; il est cartographique. L'autre, sonore, est une sorte de synthétiseur complexe dont la production de son est conditionnée par les densités de présence (la distribution) des particules. La composition sonore est en quelque sorte une conséquence décalée, indirecte, du mouvement des particules et donc des flux météorologiques. Elle est une émanation sonifiée réagissant non pas systématiquement, mais par le biais de sa propre logique. Elle n'est pas directement contrôlée, conduite par les données, mais influencée. Cette idée

d'influence, contre celle de contrôle, est centrale dans ce travail<sup>305</sup>. Il y a coexistence et influence d'un système dont le comportement est « naturel » (au sens de simulation, où il traduit directement, selon des règles visuelles explicites, le mouvement de flux) avec un autre qui est une construction conceptuelle et esthétique.

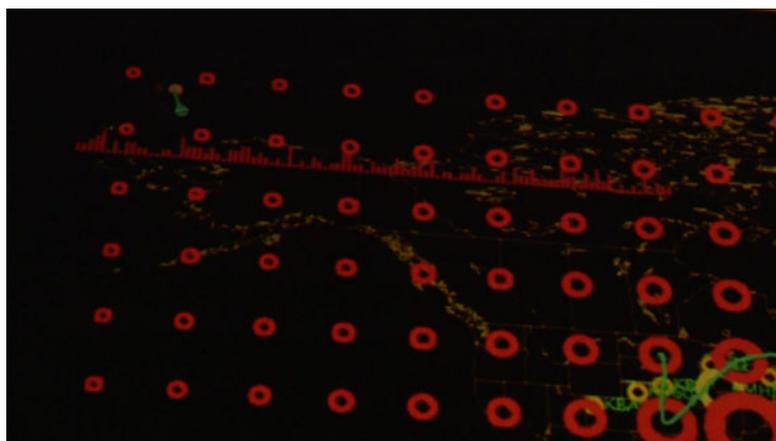


Fig. 78 : Peter Beyls, *WindChime* (2016-2017)  
© Université catholique de Porto

Dans de nombreuses pièces numériques, une correspondance directe de paramètres est mise en œuvre. Un paramètre visuel de mouvement est ainsi directement lié à un paramètre sonore, comme dans le Theremin. Il s'agit en général d'un « simple » instrument, qui traduit et met en œuvre des mappages systématiques.

Peter Beyls parle lui de recherche de la perception intuitive d'une corrélation. Celle-ci nécessite de générer, d'écouter et regarder, de retravailler, à la manière d'une sculpture, afin de trouver les paramètres qui donneront cette sensation.

---

305 Peter Beyls, « WindChime: database driven mapping Of Complex dynamical systems », *op. cit.*

## Deuxième partie.

### Dramaturgies interactives, dialectique des images et des sons dans les média numériques

Dans sa préface à l'ouvrage de Jussi Parikka consacré à l'archéologie des média<sup>306</sup>, Emmanuel Guez — s'appuyant sur la méthodologie du théoricien Friedrich Kittler — souligne que les études médiatiques permettent d'étudier ce qui se produit à l'intérieur des machines :

[...] Leur côté obscur en quelque sorte. Un art des profondeurs ou une descente aux enfers, que Kittler a identifiés en montrant que la manipulation de l'axe-temps était l'élément le plus essentiel des médias. Les données enregistrées et stockées peuvent être artistiquement manipulées, le son comme l'image pouvant être coupés, collés, montés, lus à l'endroit comme à l'envers [...], altérés tout en demeurant lisibles (voir l'art du *glitch*)<sup>307</sup>.

En effet, avec les techniques de fixation des sons et des images du cinéma émergent corollairement des esthétiques de la manipulation du temps et du montage. Ce qui est saisi (au sens de capturé) du réel est recomposé, délinéarisé. Renversements, détournements, ressaisies, la visée des œuvres ici est de fabriquer un nouveau réel, d'opérer avec lui une distance. La synchronisation des images et des sons stabilisée à partir des années 1930<sup>308</sup>, couplée au développement de la musique à l'image entraîne une réflexion sur les relations que peuvent entretenir les images animées avec les sons. Musique, parole et bruitages deviennent peu à peu un matériau plastique. Les sons et les

306 Jussi Parikka, *Qu'est-ce que l'archéologie des média ?*, Saint-Martin-d'Hères, UGA Éditions, coll. « Savoirs littéraires et imaginaires scientifiques », 2017.

307 Emmanuel Guez, « Préface », in Jussi Parikka, *op. cit.*, p. 24.

308 Notons que les recherches sur la synchronisation entre image et son ont démarré dès les débuts du cinéma. Citons notamment le kinétophone d'Edison, un caisson contenant un phonographe qui se déclenche en même temps qu'un kinétoscope, un dispositif de projection réduite aussi inventé par Edison. Il s'agit d'un système à visionnage et à écoute individuels. L'utilisateur du kinétophone était muni d'écouteurs acoustiques (identiques au stéthoscope inventé depuis le milieu du XIX<sup>e</sup> siècle) qui propagent le son du phonographe à cylindre. Il semble que plusieurs films sonores aient été réalisés, dans le cadre du Black Maria (le studio d'Edison), mais n'a été conservé que le premier, *The Dickson Experimental Sound Film* (1895). Celui-ci a été réalisé par William K. L. Dickson, l'assistant d'Edison, qui y fait aussi une apparition. (pour d'autres informations, voir par exemple Laurent Mannoni, *Le grand art de la lumière et de l'ombre. Archéologie du cinéma*, Paris, Nathan, 1994, p. 374-375)

images qui étaient saisies dans un même temps, par deux appareils (synchronisés après coup grâce au clap, ou bien à l'aide d'un câble<sup>309</sup>) vont progressivement laisser la place à des caméras légères et synchrones<sup>310</sup>, vidéo ou des appareils numériques capables de saisir l'image et le son simultanément. Parce que la synchronisation audiovisuelle est constitutive du processus de saisie à l'ère numérique, la désynchronisation devient en contrepartie un choix esthétique<sup>311</sup>.

Cette partie se consacre ainsi aux concepts et aux processus mis en œuvre dans les formes interactives héritées de ce cinéma de synchronisation. C'est ce que Jean-Louis Boissier nomme la dramaturgie de l'interactivité<sup>312</sup>. Nous examinerons, dans un premier chapitre, comment les procédés de l'inversion et du ralenti — qui existent depuis les débuts de l'enregistrement du son, puis, ultérieurement, celui de l'image animée — ont fait évoluer la relation entre image et son dans le contexte des arts interactifs. Dans un deuxième chapitre, nous nous attarderons sur le saut interactif. Procédé fondamental, il autorise une navigation interne et instantanée dans un même segment audiovisuel, autrement dit : la téléportation d'un moment à un autre. Devenant contrôlables par le programme informatique, le saut et la saccade, utilisés frénétiquement, vont permettre aux artistes d'étirer ou de disloquer la microstructure du fragment. Prolongement conceptuel du saut, l'accès instantané permis par le numérique a fait naître de nouvelles formes artistiques. Celles-ci, fondées sur des bases de données, font interagir utilisateurs, interfaces et programmes. Elles constitueront l'objet du troisième chapitre. Le réemploi et la performativité, pratiques démocratisées par les Djs et des VJs y seront également explorées. Nous nous intéresserons enfin, dans un

309 Dans les années 1950, des dispositifs de tournage synchrone en extérieur (sans recours au secteur) voient le jour. Un câble, souvent qualifié de « fil à la patte », reliait l'opérateur son et l'opérateur image. Celui-ci permettait une capture synchrone, mais restreignait leurs possibilités de mouvement (Vincent Sorrel, « L'invention de la caméra Éclair 16 : du direct au synchrone », *Revue 1895* n° 82, 2017, <https://journals-openedition-org.ezpaarse.univ-paris1.fr/1895/5330>, consulté le 30/01/23).

310 Par exemple, les caméras Éclair, utilisées par les documentaristes et les journalistes. Elles étaient capables de capturer l'image et le son en synchronisation, et d'impressionner directement la pellicule du film avec les images et la bande-son.

311 C'est l'hypothèse qu'émet Vincent Bouchard pour le médium vidéo : « L'asservissement de l'image et du son étant une donnée "ontologique" du médium vidéo, la désynchronisation devient un choix esthétique produisant un changement dans la perception de la temporalité des images et des sons. » (Vincent Bouchard, « Du Nagra au caméscope : questions de synchronisation image/son », *Intermédialités* n° 19, printemps 2012, « Synchroniser », 2012, <https://www.erudit.org/en/journals/im/2012-n19-im0305/1012659ar/>, consulté le 30/01/23).

312 Jean-Louis Boissier, « Dramaturgie de l'interactivité », in Jean-Louis Boissier, *La relation comme forme. L'interactivité en art*, Genève, MAMCO, 2004, p. 23-29.

quatrième et dernier chapitre, aux médias audiovisuels situés. Entre voyage imaginaire, surveillance et matérialité du numérique, nous explorerons des œuvres qui fabriquent de nouvelles modalités de saisie, par l'image et le son, de nos environnements naturels et urbains.

## **2.1 Programmer l'axe-temps des images-son : inverser, ralentir, accélérer**

### **2.1.1 Archéologie des sons et des images animées inversés**

Aussi bien dans l'appréhension du rôle du numérique dans les arts que dans la compréhension des relations esthétiques entre les images et les sons, le travail du théoricien Friedrich Kittler autour de l'archéologie des médias nous livre un éclairage essentiel. Il pointe en particulier la période de la fin du XIX<sup>e</sup> siècle, à laquelle des appareils, pour la première fois, parviennent à enregistrer le réel, par la fixation de l'audible et du visible.

Ce que rendirent d'abord enregistrable le phonographe et le cinématographe, dont les noms ne dérivent pas de l'écriture par simple coïncidence, fut le temps : sous la forme de mélanges de fréquences des bruits pour l'acoustique, sous celle d'une suite d'images en mouvement pour l'optique. Le temps constitue la limite de tous les arts. Il faut d'abord immobiliser le flux de données du quotidien, avant que celui-ci puisse devenir image ou signe. Ce qu'on appelle le style en art n'est rien d'autre que le mécanisme de commutation de ces échantillonnages et de ces sélections<sup>313</sup>.

En ce qui concerne l'audible, l'inscription de vibrations sonores sur les sillons des disques du phonographe, en 1877, est l'un de ces moments clés. Même si l'invention d'Edison résulte de nombreuses découvertes (parmi elles celles de Léon-Scott de Martinville, puis de Charles Cros), elle pose un jalon important car elle est, selon Friedrich Kittler, la première machine capable de « d'écrire et de lire, stocker et scanner, enregistrer et jouer<sup>314</sup> », capacités techniques que l'on retrouve également dans les média informatiques. Grâce au phonographe d'Edison en effet, il était à la fois possible

---

<sup>313</sup> Friedrich Kittler, *Gramophone, Film, Typewriter*, Dijon, Les presses du réel, 2018, p. 38-39.

<sup>314</sup> *Ibid.*, p. 77.

d'enregistrer et de jouer, mais aussi d'accélérer, de ralentir, ou de changer le sens de lecture. Ces possibilités de variation offrent de nouvelles perspectives esthétiques.

Pour le cinéaste et théoricien Jean Epstein, il existe deux niveaux de défilement ; Paul-Emmanuel Odin, dans son ouvrage *L'inversion temporelle au cinéma*, en résume les sens ainsi :

Le ralenti et l'accélééré sont les axes de « variabilité du temps » (au niveau quantitatif), alors que l'inversion temporelle est sur l'axe de la « réversibilité » (la qualité du temps est affectée). Cette différence nous importe au plus haut point car le qualitatif, opposé au quantitatif, est de l'ordre du sens, d'un supplément de signification<sup>315</sup>.

Dans le registre de l'image, l'inversion du sens de lecture était déjà possible avec les jouets optiques du précinéma tels que le zootrope. En 1895, les frères Lumière affirment cette réversibilité avec le cinéma, en projetant des films comme *Démolition d'un mur* ou *Les bains de Diane* à l'envers, à la fin des séances<sup>316</sup>. Les premiers sons inversés sont produits quant à eux par Edison avec son phonographe. L'effet semble notamment impressionnant avec les sons de percussion<sup>317</sup>, dont le timbre devient extrêmement étrange lorsque ceux-ci sont joués à l'envers. Philippe Langlois, dans son ouvrage *Les cloches d'Atlantis*<sup>318</sup> rappelle que c'est la fixation du son sur le support linéaire qu'est la piste optique de la pellicule, qui, dans les années 1930, a induit une potentialité d'action sur la vitesse et le sens de lecture de la piste sonore du film.

Cela implique que la lecture optique, tout comme la bande magnétique vingt ans plus tard, peut être lue dans les deux sens. Il résulte de cette inversion une permutation des phénomènes acoustiques : les résonances précèdent les attaques, les nuances s'inversent, les mélodies sont lues à l'envers. Pour chaque son entendu, au lieu du cheminement naturel : émission, résonance, extinction, il s'opère un renversement total de ces paramètres. C'est ainsi que, passé à l'envers, un coup de mailloche sur une cymbale donne l'impression que le son émerge progressivement du silence, pour ensuite croître en intensité jusqu'à

---

315 Paul-Emmanuel Odin, *L'inversion temporelle au cinéma*, Dijon, Les presses du réel, coll « Al Dante », 2014, p. 37.

316 *Ibid.*, p. 44.

317 *Ibid.*, p. 65.

318 Consacré aux liens entre musique électroacoustique et cinéma.

son point culminant avant d'être brusquement aspiré et disparaître dans le silence qui a précédé son émission<sup>319</sup>.

Plus encore, Langlois décrit un procédé tout à fait nouveau utilisé dans ce contexte du son optique des années 1930 : la partition rétrograde. Il s'agit pour le compositeur de transcrire sa composition sous la forme d'une partition traditionnelle, puis de la réécrire, dans la chronologie inverse : la dernière note devient la première, la première la dernière. L'orchestre joue la partition inversée, puis la bande est retournée après l'enregistrement. L'ordre originel est retrouvé, mais la résonance de chaque son précède son attaque. La partition rétrograde a été utilisée dans des films comme *Rapt* (1933) de Dimitri Kirsanoff, sur la musique d'Arthur Hoérée et d'Arthur Honegger, ou dans *Zéro de conduite* (1933), de Jean Vigo sur la musique de Maurice Jaubert. Ces techniques ont aussi été très utilisées avec Pierre Schaeffer et la musique concrète dans les années 1950. Parmi les utilisations contemporaines, dans les années 1990, David utilise ce procédé dans une scène clé du film *Twin Peaks Fire Walk with me* (1992), qui se déroule dans la chambre rouge (la Black Lodge). Les acteurs ont appris leurs dialogues grâce à une transcription phonétique du texte lu à l'envers, et ont effectué leurs déplacements en arrière lors du tournage<sup>320</sup>. Tout a été inversé au montage, créant un temps étrange et décalé. La rétrogradation et l'inversion servent souvent chez le réalisateur à créer un trouble, une étrangeté. Dans le monde du religieux, lire à l'envers des paroles sacrées constitue une invocation au diable, aussi ces techniques de *Backmasking*<sup>321</sup> (l'enregistrement de musiques ou de paroles à l'envers), utilisée pourtant par des groupes comme les Beatles, sont associées, pour beaucoup, aux complots et messages cachés. Les ouvriers des usines Lumière, voyant les films à l'envers, parlaient de leurs patrons comme des sorciers<sup>322</sup>. Certaines techniques sonores, comme le son delta, consistent à « créer un son nouveau en collant à un son donné sa réplique inversée temporellement<sup>323</sup> », c'est-à-dire à superposer un son et son double inversé.

---

319 Philippe Langlois, *Les cloches d'Atlantis : musique électroacoustique et cinéma, archéologie et histoire d'un art sonore, op. cit.*, p. 179.

320 *Ibid.*, p. 191.

321 *Ibid.*, p. 191.

322 Paul-Emmanuel Odin, *op. cit.*, p. 46.

323 *Ibid.*, p. 69.

L'endroit et l'envers sont comme les faces d'une même pièce [...]. Il suffit d'un tour de manivelle et le temps se replie sur lui-même, en une sorte d'anneau de Moebius où le couple endroit-envers s'annule, comme si le cinématographe devait s'évanouir pour faire comme s'il ne s'était rien passé, pour revenir à l'origine du temps qui en même temps s'indique comme la fin du temps<sup>324</sup>.

### 2.1.2 La *Time Machine* de Piotr Kowalski. Renversements temporels de l'image et du son.



Fig. 79 : Piotr Kowalski, *Time Machine II* (1981)  
© Piotr Kowalski / piotrkowalski.com

Grâce aux possibilités — même encore limitées — de stockage et de programmation des premiers micro-ordinateurs, ces expérimentations d'inversion vont pouvoir s'opérer en direct. Il devient en effet possible de découper et de recomposer en temps réel les flux issus d'une captation. Ce procédé est utilisé de manière pionnière par l'artiste Piotr Kowalski, qui va chercher à inverser l'écoulement naturel du temps saisi grâce à une machine informatique : la *Time Machine*. Deux versions seront développées : l'une, seulement sonore, en 1979 ; la seconde, audio-visuelle, en 1981.

---

324 *Ibid.*, p. 47.

### ***Un Miroir sonore de la voix : la Time Machine I (1979)***

L'ambition de Piotr Kowalski, avec sa première *Time machine*, est de réaliser un dispositif de miroir sonore capable d'inverser, en temps réel, un flux capté en direct par un microphone. Ce renversement, qui est déjà rendu possible techniquement par la manipulation d'une bande magnétique, ne l'est toutefois qu'après enregistrement. Entre la France et la Suède durant l'année 1970, Kowalski réussit à développer, en détournant un magnétophone à têtes tournantes, un système fonctionnel. Un encodage linéaire du signal et un décodage étaient effectués par des têtes tournantes, qui tournaient à une vitesse deux fois plus rapide que la bande, ce qui produisait un découpage des séquences en morceaux ; chaque morceau était lu à l'envers. Ce système s'avéra finalement d'une fragilité trop critique pour le rendre utilisable. Pour pouvoir comme il le souhaite inscrire un signal dans une mémoire, puis le restituer, il lui faudrait utiliser un microprocesseur, mais ces technologies ont un coût encore trop élevé. Kowalski va devoir attendre la fin des années 70, avec son arrivée au Center for Advanced Visual Studies de Cambridge, au sein de l'Architectural machine Group du M.I.T., pour pouvoir intégrer les technologies informatiques à son projet.



Fig. 80 : Piotr Kowalski, *Time Machine I* et installation galerie Feldman à New York (1979)  
© Centre Georges Pompidou/Musée national d'art moderne

La *Time Machine I* est présentée pour la première fois à New York entre mai et juin 1979, à la galerie Ronald Feldman. Elle est constituée d'une valise en aluminium contenant le dispositif et d'un boîtier de commande sous plexiglas, complétée d'un

microphone et d'un casque. Elle forme, avec une autre œuvre intitulée *Miroir*, un ensemble. Cette sculpture propose un système de renversement, visuel cette fois-ci, qui, grâce à un système constitué d'un socle tournant et de miroirs, donne à voir au visiteur son image non inversée. Kowalski décrit le fonctionnement de sa machine ainsi :

Techniquement la *Time Machine I* est un ordinateur construit et programmé afin de renverser la direction du temps perceptible, telle qu'elle est mesurée par le son. Elle travaille en temps réel et convertit le signal sonore en donnée numérique, stocke celle-ci puis l'inverse et la restitue à son tour en un signal sonore qui sera donc inversé par rapport au premier. [...] La *Time Machine I* coupe le cours du temps en segments discrets dont la longueur est choisie par l'opérateur (par le public). Chaque segment est travaillé par la machine, un à un. Quand un segment est lu à l'envers par la mémoire, un autre est stocké et transformé en donnée numérique pour la lecture suivante<sup>325</sup>.

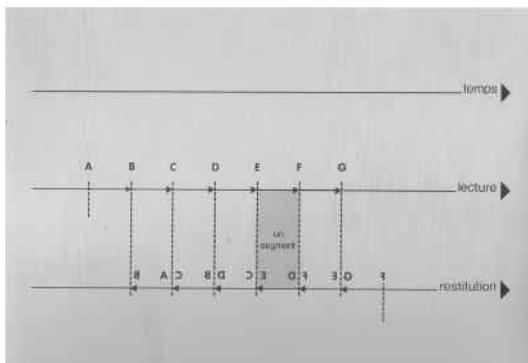


Fig. 81 : Piotr Kowalski, Schéma du principe des *Time Machine*  
© Centre Georges Pompidou/Musée national d'art moderne

Les segments découpés peuvent durer de 1/1000 e de seconde à 1 seconde. Un décalage, correspondant à cette durée d'échantillonnage, existe donc entre le premier segment et son double inversé.

### **Poétique de la voix manipulée**

Il n'est pas anodin que la voix et le magnétophone soient à l'origine du projet de *Time Machine*. Dans différents textes, Kowalski fait référence à l'histoire des différents ins-

<sup>325</sup> Piotr Kowalski. *Time Machine et projets*, catalogue d'exposition, Paris, Centre Georges Pompidou, Musée national d'art moderne, 1981, p. 45.

truments d'enregistrement, du cylindre de cire qui pouvait capturer la voix, jusqu'aux films magnétiques. Kowalski utilisait également comme matériau d'expérimentation des lectures de William Burroughs ou d'autres poètes (parfois en repassant des extraits inversés pour les remettre à l'endroit). Il faut rappeler concernant la voix que dans les années 1950, le magnétophone fut une révolution pour les poètes, qui purent désormais inscrire facilement leur voix sur un support grâce à un matériel léger utilisable chez eux, donnant naissance à ce que l'on a appelé la poésie sonore. L'un de ces pionniers, Bernard Heidsieck, après avoir écouté Stockhausen, comprit que la poésie avait des années de retard sur les musiques nouvelles et commença à l'utiliser dans ses performances (par exemple les célèbres poèmes-partitions) en mettant en scène la simultanéité de sa propre voix enregistrée avec sa voix directe, performée. Le poète Henri Chopin fut quant à lui l'un des premiers à travailler sur la manipulation temporelle de bandes vocales et sur leur superposition. L'œuvre de poésie sonore qui pourrait tout-fois se rapprocher le plus des concepts de la *Time machine* serait probablement *I Am That I Am* (1958) de Brion Gysin, travail précurseur de l'art génératif textuel informatique, et inspiré par une tautologie d'Aldous Huxley<sup>326</sup>. D'une part Gysin y opère toutes les permutations possibles d'une même phrase, dans une sorte de cut-up (qu'il a lui-même inventé avec Burroughs) de son propre texte ; il discrétise un texte par le mot et le réassemble. D'autre part, il travaille des effets temporels sur la bande magnétique. Kowalski, avec sa *Time Machine* veut en quelque sorte appliquer ce type de procédé très algorithmique, mais sur une captation en direct et en réassemblant les segments à l'envers<sup>327</sup>.

### ***La Time Machine II (1981). Du son à l'image-son***

En 1981 est présentée au Musée national d'art moderne — Centre Pompidou une nouvelle version de la *Time Machine*, la *Time Machine II*. Celle-ci, s'appuyant sur le même

<sup>326</sup> Interview de Brion Gysin par Gérard-Georges Lemaire en 1974, in Henri Chopin, *Poésie sonore internationale*, Paris, Jean-Michel Place, 1979, p. 128.

<sup>327</sup> Ce ne sera pas le dernier travail de Kowalski sur la voix, il sera à nouveau précurseur dans l'utilisation du numérique (l'impression 3D notamment) en transcrivant dans une sculpture intitulée *Passionnement* (1999) un sonogramme de la voix du poète Gherasim Luca prononçant ce même mot. La sculpture, constituée de 99 lames de verre, a nécessité l'utilisation de techniques liées aux empreintes vocales et à la paramétrisation de données (*Laboratoires de l'art*, catalogue d'exposition, Luxembourg, Paris, Mudam Luxembourg, Hermann, Musée des arts et métiers-le Cnam, 2016, p. 79).

procédé, intègre l'image. Ne seront plus renversés des flux sonores, mais des flux audio-visuels, des segments d'image-son dont le couplage initial est conservé.

Le dispositif, plus complexe et plus étoffé, est toujours présenté avec la pièce *Miroir*. Deux moniteurs vidéo, avec diffusion sonore, sont positionnés dans l'espace de l'installation. Celui de gauche renvoie le flux audiovisuel direct, celui de droite le flux audiovisuel inversé. Dans un grand cube transparent, se trouve la machine, montrant au visiteur la matérialité de la machine. Un pupitre de commande permet au visiteur de choisir l'une des 4 entrées vidéo (couplées à l'audio correspondant), ambiantes ou diffusées.



Fig. 82 : Piotr Kowalski, *Time machine II* (1981)  
© Piotr Kowalski / piotrkowalski.com

La vidéo sonore peut être celle d'un signal de télévision, d'une caméra pointée sur la piazza du Centre Pompidou, d'une caméra avec téléobjectif pointée vers l'intérieur, ou bien d'une caméra pointée sur l'espace de l'installation (avec microphones). Il est possible de choisir la fréquence de découpage, et d'effectuer des arrêts sur image. Les temps de renversement vont de 0 seconde à 3 secondes et demie<sup>328</sup>. Entre les deux moniteurs, un ensemble de deux microphones et d'une caméra permet une captation audio-vidéo en temps réel du visiteur, s'il le souhaite.

<sup>328</sup> Commentaire audio de Piotr Kowalski, dans le documentaire *Deux temps, trois mouvements*, de Gisèle et Luc Meichler, 1982. Visible en ligne : partie 1, <https://www.dailymotion.com/video/x411g8> et partie 2, <https://www.dailymotion.com/video/x411g8>, consultés le 30/01/23. La nature du documentaire est intéressante, car ce dernier utilise dans son montage même la *Time Machine*, créant une mise en abyme de la pièce.

Jean-Christophe Bailly décrit le travail de Kowalski comme un détournement d'outils provenant de la science et des techniques, à la fois critique et romantique. Le discours de Kowalski sur les techniques est à ce titre très clair :

Mais les nouvelles technologies nous l'imposent car sans ça on les subit. Pour ne pas les subir, il faut d'un côté les simplifier, les mettre à plat, et de l'autre permettre le déchiffrement de leur fonctionnement. C'est pour apprendre à ne pas subir. Si on peut jouer avec quelque chose, d'abord on ne le prend pas au sérieux et deuxièmement on voit qu'on peut le changer. C'est le début de l'apprentissage de la liberté ou du traitement démocratique du monde. Une constante dans mes préoccupations est de produire des objets qui servent à démontrer que l'on peut devenir acteur du réel, que les choses ne sont pas figées, qu'on peut les modifier à chaque moment, que rien n'est déjà là, donné<sup>329</sup>.

En inversant des flux audiovisuels en direct, Kowalski montre que l'utilisation du numérique peut produire des expériences interactives nouvelles. Dans celles-ci, la voix et l'image du visiteur, ou bien des saisies du réel, se confrontent en miroir à leurs répliques manipulées.

### 2.1.3 Pratiques fondatrices du ralenti

Comme Emmanuel Guez l'a relevé<sup>330</sup>, Friedrich Kittler a insisté sur l'importance, dans les médias techniques, de la manipulation de l'axe-temps. La vitesse et le sens de lecture sont en quelque sorte liés depuis la fixation des images et des sons. Dès l'invention du phonographe par Edison<sup>331</sup>, celui-ci a pu, régulièrement, lors de présentations de sa machine, tourner la manivelle plus vite, ou plus lentement, afin de moduler les morceaux pour amuser le public. Pour Kittler, la portée de cette action est toutefois plus large :

Lorsque la vitesse de lecture du phonographe diffère de celle de l'enregistrement, ce ne sont pas seulement les sons clairement audibles qui varient, mais les spectres des bruits tout entiers. Ce qui est manipulé est le réel, et non plus le symbolique<sup>332</sup>.

329 Propos recueillis par Jean-Louis Boissier et Jean-Marie Dallet, Montrouge, le 24 septembre 1992, [http://www.ciren.org/artifice/artifices\\_2/kowalski.html](http://www.ciren.org/artifice/artifices_2/kowalski.html), consulté le 30/01/23.

330 Voir l'introduction à cette partie.

331 Comme beaucoup d'inventions, les principes d'inscription préexistaient à Edison, notamment avec Charles Cros ; mais c'est bien Edison qui a pu rendre la machine exploitable.

332 Friedrich Kittler, *op. cit.*, p. 82.

Par cet exemple, il illustre ce qu'il a décrit comme le passage du symbolique au réel, grâce à la découverte par les physiologistes et les scientifiques de la fréquence sonore. Les voix, la musique et les bruits sont traités à rang égal et fusionnent dans leur capture. Dans son commentaire du film *Zéro de conduite* de Vigo, Philippe Langlois pointe la complémentarité audiovisuelle dans certaines scènes du film, car des jeux s'opèrent entre le ralenti visuel (qui nécessite un tournage en accéléré) et la rétrogradation sonore<sup>333</sup>. Pour Jean Epstein, le ralenti est synonyme de mortification et de matérialisation.

Plus ralentie encore, toute substance vive retourne à sa viscosité fondamentale, laisse monter à sa surface sa nature colloïdale foncière. Enfin, quand il n'y a plus de mouvement visible dans un temps suffisamment étiré, l'homme devient statue, le vivant se confond avec l'inerte, l'univers évolue en un désert de matière pure, sans trace d'esprit<sup>334</sup>.

Ainsi pour lui, si l'on freine le cours du temps et donc le mouvement des êtres, « on en fait disparaître ou on en détruit la qualité vitale<sup>335</sup> ». Le philosophe Elie During, dans un article consacré à l'effet Bullet Time<sup>336</sup>, définit trois fonctions du ralenti. Une fonction emphatique, c'est-à-dire « sa capacité à suggérer la puissance et la vitesse d'un geste, ou à magnifier la virtuosité de celui qui l'accomplit<sup>337</sup> ». Une fonction d'amplification ou d'hyperbole, qui a pour but d'intensifier « la dimension dramatique ou expressive d'une scène, ou bien en surligner la gravité, la solennité<sup>338</sup> ». Enfin, il peut remplir une fonction lyrique ou onirique, un « corps suspendu dans sa chute, ravi à lui-même, a quelque chose des images flottantes du rêve<sup>339</sup> ». Nicole Brenez, spécialiste du cinéma d'avant-garde et expérimental, montre dans un article consacré à ces questions combien les cinéastes ont pu domestiquer la vitesse, jusqu'à utiliser le plan — se-

333 Philippe Langlois, *op. cit.*, p. 187.

334 Jean Epstein, *L'Intelligence d'une machine*, Paris, Jacques Melot, 1946, p. 23, version numérique, [http://classiques.uqac.ca/classiques/epstein\\_jean/intelligence\\_machine/intelligence\\_une\\_machine.pdf](http://classiques.uqac.ca/classiques/epstein_jean/intelligence_machine/intelligence_une_machine.pdf), consulté le 30/01/23.

335 *Ibid.*, p. 24.

336 Effet ralenti extrême, avec rotation du point de vue de la caméra. Démocratisé en 1999 avec le film *Matrix*.

337 Elie During, « L'image ralentie : de la caméra-œil au bullet-time », in Jérôme Game (dir.), *Images des corps/Corps des images au cinéma*, Lyon, ENS Éditions, coll. « Signes », 2010, <https://books.openedition.org/enseditions/9081?format=toc>, consulté le 30/01/23.

338 *Ibid.*

339 *Ibid.*

lon ses mots — comme boîte de vitesses<sup>340</sup>, car susceptible de changer de rythme dans un même mouvement. Dans le cas du ralenti, soit à travers l'enrichissement de leur palette, ou bien, de manière plus radicale, dans un approfondissement thématique. C'est le cas dans *See You Later/Au revoir* (1990), de Michael Snow. Il s'agit d'un seul plan-séquence de 30 secondes, dans lequel un employé part de son bureau, prend son manteau, prononce la phrase « au revoir », et sort de la pièce. Snow dilate le temps de cette scène en 17 minutes, transformant un moment du quotidien en expérience éprouvante pour nos sens, peu habitués à un ralenti de cette durée. Le ralenti peut aussi provenir d'un « refilmage d'œuvres, métamorphosées par leur transfert dans un autre médium<sup>341</sup> », comme *24 Hours Psycho* de Douglas Gordon (1993). Le film original d'Alfred Hitchcock y est passé en 2 images par seconde, étirant la durée du film à 24 heures. Il faudrait également citer Bill Viola, qui utilise le ralenti visuel dans bon nombre de ses œuvres vidéo.

#### 2.1.4 *The Waves*, de Thierry Kuntzel. Passages du mouvant à l'immobile, du sonore au silence



Fig. 83 : Thierry Kuntzel, *The Waves* (2003)  
© Galerie de l'ésam, site de Caen

340 Nicole Brenez, « Flux, flash et prisme. Vitesses critiques dans l'histoire du cinéma », in *Le Genre humain* n° 49, 2010, <https://www-cairn-info.ezpaarse.univ-paris1.fr/revue-le-genre-humain-2010-1-page-125.htm>, consulté le 30/01/23.

341 *Ibid.*

La *Time Machine* de Kowalski, dans une section précédente de ce chapitre, nous a donné l'exemple d'une rétrogradation temporelle qui serait impossible à réaliser sans l'ordinateur. Les microfragments audiovisuels provenant d'une saisie en temps réel sont rediffusés dans l'ordre inverse de leur capture, et ce quelle que soit la nature de son flux audiovisuel. Dans l'installation interactive *The Waves* (2003)<sup>342</sup> de Thierry Kuntzel, que nous allons explorer dans ce chapitre, la microstructure n'est pas affectée. Le visiteur est en revanche acteur d'une interactivité qui va le conduire à éprouver des passages de l'animé à l'immobile, du son au silence, par l'intermédiaire d'une seule manipulation temporelle : le ralenti.

### ***Boucles et va-et-vient***

Dans le fond d'un espace plongé dans l'obscurité, le visiteur peut observer un écran sur lequel se déroule en boucle une séquence vidéo lui présentant des vagues s'abattant sur le rivage, sans aucun indice de leur localisation. Le bruit des vagues, a priori le son direct capté dans le même temps que l'image, est diffusé sur des haut-parleurs le long du parcours. Au fur et à mesure que le visiteur avance, il constate que l'image animée ralentit, qu'elle perd progressivement ses couleurs, et que le son semble se modifier, se ralentir. À un certain point, le ralentissement est tel que l'image se retrouve figée, en noir et blanc, dans un espace devenu silencieux. Si l'on recule, l'image et le son reprennent doucement vie, mais il faut, pour retrouver l'énergie originale, revenir à son point de départ.

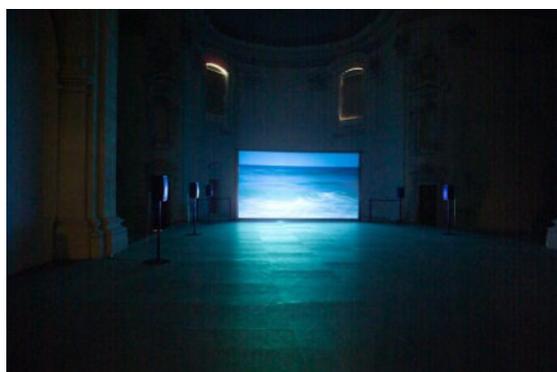


Fig. 84 : Thierry Kuntzel, *The Waves* (2003)



Fig. 85 : Thierry Kuntzel, *The Waves* (2003)

<sup>342</sup> Une version est disponible en ligne sur le YouTube de La Compagnie, <https://www.youtube.com/watch?v=LL5czzhM0Q>, consulté le 30/01/23.

Il y a trois niveaux de va-et-vient dans *The Waves*. En premier lieu, la séquence des vagues, contemplative, est une boucle : quels que soient les déplacements qui se produisent, lorsque la fin, quasi indicible, est atteinte, la séquence est rejouée. Ensuite, on peut considérer les mouvements de mer comme un ensemble « génératif », constitué d'instances du même élément « souche », la vague. Chaque vague porte sa propre individualité, ses propres qualités de force ou de hauteur, mais de fait appartient à la même espèce. Enfin, comme le décrit Thierry Kuntzel, quand le spectateur recule, le même revient : « il y a retour du même, ou du faux même<sup>343</sup> ». Le public fabrique, par ses propres mouvements, des va-et-vient physiques dans l'espace de l'installation.

### ***Soudure perceptive et mappage de paramètres***

Il est important de noter que l'image et le son (ou le son et son image), restent constamment synchronisés dans le ralenti. Dans sa thèse, le chercheur et artiste Vincent Ciciliato qualifie ce procédé de soudure perceptive.

Ce principe de variabilité est l'un des garants pour la cohésion perceptive de l'objet en tant qu'unité insécable. L'image de la mer et de son bruit sont ici perçus comme une seule et même matière, et l'affection de l'une induit automatiquement une altération de l'autre. C'est le code temporel de la bande vidéo, ainsi que sa modulation par le spectateur, qui servira de point de jonction entre les niveaux sensitifs. En agissant sur cette temporalisation commune du son et de l'image on affecte par la même occasion leur exécution. Mais étrangement, c'est cette même mise en mouvement et altération parallèle des deux modalités sensorielles, qui en assure une réelle jonction, et leur soudure perceptive<sup>344</sup>.

En termes techniques, une même transformation est affectée à l'image et au son. À partir du moment où le visiteur pénètre dans l'espace de l'installation, le dispositif numérique va créer une liaison entre le déplacement de son corps et la manipulation de

---

343 Interview de Thierry Kuntzel dans *L'eau*, documentaire de Jean-Yves Jouannais, <https://vimeo.com/58473589>, consulté le 30/01/23.

344 Thèse de Vincent Ciciliato, *Glitch(s) imago-sonore(s) : du processus d'objectivation par découpe et synchronisation son-image à une esthétique du micro-mouvement*, Amiens, 2010, sous la direction de François Parfait, p. 327.

l'axe-temps de la vidéo. Ceci est produit assez simplement, grâce à un capteur de distance (ici une caméra) qui va permettre *le mapping* de la distance du corps à l'écran sur trois paramètres de la vidéo : la vitesse de défilement de la vidéo (son *framerate*), la vitesse de lecture du son, et le degré de saturation des couleurs qui devient de plus en plus faible lorsque le spectateur se rapproche de l'écran.

Les imbrications permanentes de « mouvements » (celui de la vidéo, celui des vagues, celui du participant), lorsque le visiteur arrive près de l'écran, donnent lieu à une solennité particulière : l'ensemble se fige, le son n'est plus audible. Est-il stoppé par programmation, ou bien ne peut-il tout simplement pas se développer dans ce ralenti extrême<sup>345</sup> ? Quelle qu'en soit la raison, comme le rappelle Michel Chion : « Il n'y a pas d'arrêt sur le son<sup>346</sup> ». Les perceptions des objets visuels et des objets sonores dans le contexte de la *pause*, sont différentes. La nature continue du sonore appelle cette radicalité. Si l'on met un son en pause, il y a silence, brutal, et disparition. Ou plus précisément, suspension.

### ***L'absence du son***

L'expérience de l'installation mêle, à notre sens, deux sentiments contradictoires. D'un côté la satisfaction de contrôler cet ensemble visuel et sonore, mais en même temps la frustration de devoir rester éloigné pour éprouver le maximum de mouvement, de couleur et de son. Comme le note tout à fait justement Caroline Chik, dans son ouvrage dédié aux passages entre l'image fixe et l'image animée, cette frustration réside également dans le fait que le point de passage au fixe se situe tellement près de l'écran que l'on ne voit plus l'image, même figée, dans son ensemble ; nous n'en voyons plus que la dimension numérique, les points :

En effet, paradoxalement, le spectateur ne verra jamais la vague qu'il gèle lui-même, l'arrêt se produisant seulement lorsqu'il est trop proche de l'écran pour le percevoir. [...] À l'analyse du mouvement en tant qu'arrêt du temps s'est substituée l'analyse de l'image-écran en tant qu'espace. En intervenant au niveau de la matière temporelle de l'image, ce n'est plus le photogramme qui res-

345 Il est intéressant de noter que techniquement, il n'est normalement pas nécessaire de programmer cette disparition du son, c'est ce qu'il se passera informatiquement sur toute vidéo ralentie par programmation, qui dépasse le seuil de la cadence des images. L'artiste peut toutefois aussi avoir programmé le silence à un certain seuil d'approche.

346 Michel Chion, *Le son*, Paris, Nathan, 2000, p. 151.

surgit à travers l'arrêt, mais la matière spatiale, lumineuse, électronique de l'image. *The Waves* symbolise ainsi la mise en crise de l'image en mouvement favorisée par l'interactivité, au profit de la mise en valeur du signifiant, de la perte du référent, du non-voir<sup>347</sup>.

Inversé, le dispositif aurait peut-être été plus ludique pour le visiteur, qui aurait été en mesure de donner vie à la vidéo quand il s'approche. On s'approche en général de la mer pour pouvoir la toucher. Ici le bruit est donné d'emblée ; s'approcher fige, et rend silencieux, ce qui peut s'appréhender de manière assez pessimiste. Corinne Castel, productrice de la pièce, laisse entendre à ce propos que le titre *The Waves* fait bien référence au livre éponyme de Virginia Woolf<sup>348</sup>, et que l'interaction mise en jeu, par le chemin vers le silence et la fixité, fait écho à sa mort tragique. Les mots de Raymond Murray Schafer<sup>349</sup> sur le silence évoquent cet aspect.

L'homme aime produire des sons pour se rappeler qu'il n'est pas seul. De ce point de vue, le silence total constitue le rejet de la personne humaine. L'homme redoute l'absence de son, car il redoute l'absence de vie. Le silence ultime étant celui de la mort, il atteint à la dignité suprême dans la cérémonie du souvenir<sup>350</sup>.

La beauté de l'image fixe, en noir et blanc, dans le silence, nous fait réfléchir au statut du sonore face à l'image figée. Ce que l'interactivité favorise, ce sont les passages, les changements d'états, de modes. Cette scène du réel enregistré n'appelle pas de résolution. Ne se joue pas ici ce que Michel Chion nomme une vectorisation, c'est-à-dire une dramatisation des plans, une orientation vers un but<sup>351</sup>. Le son et l'image directs, flux temporels couplés même dans le ralentissement, se séparent et démontrent leur nature différente au moment de la pause. Les deux sont en suspens. Il ne s'agit pas d'une

---

347 Caroline Chik, *L'image paradoxale — Fixité et mouvement*, Villeneuve-d'Ascq, Presses universitaires du Septentrion, 2011, version numérique, <https://books.openedition.org/septentrion/46763?lang=fr>, consulté le 30/01/23.

348 Interview de Corinne Castel, productrice de l'œuvre, par Louise Mahler (2'09), <http://www.ligne16.net/le-109-en-action-pour-le-festival-movimenta/>, consulté le 30/01/23.

349 Compositeur et théoricien écologiste. Il est à l'origine dans les années 1970 de l'idée de « paysage sonore », c'est-à-dire de la capacité à appréhender nos environnements sous l'angle acoustique.

350 Raymond Murray Schafer, *Le paysage sonore, le monde comme musique*, coll. « Domaine sauvage », Paris, WildProject, 2010, p. 365.

351 Michel Chion, *L'audiovision*, op. cit., p. 16.

photographie, car nous savons qu'elle dépend d'un tout — la séquence vidéo ; il ne s'agit pas d'un silence, mais d'un son en suspens.

### 2.1.5 *Temps Libre*, de Samuel Bianchini. La voix agitatrice



Fig. 86 : Samuel Bianchini, *Temps libre* (2004)  
© Samuel Bianchini / dispotheque.org

Dans les sections précédentes, que ce soit dans la *Time Machine*, ou dans *The Waves*, le matériau utilisé est un bloc soudé et synchronisé d'image-son, que l'artiste inverse ou ralentit dans sa globalité. Dans l'installation interactive *Temps libre* (2004)<sup>552</sup> de Samuel Bianchini, il est toujours question de temps, d'un même fragment, mais l'image et le son sont affectés à des rôles tout à fait différents, dans une accélération fragmentée du temps visuel.

#### ***La revanche du sonore. L'image asservie à la voix***

L'installation consiste en la projection sur écran d'une séquence vidéo muette. On y observe, depuis un point de vue assez éloigné, un terrain de golf installé sur des falaises et donnant sur la mer. Un ou deux golfeurs cheminent dans cet espace. Le spectateur constate assez vite qu'il a à sa disposition un mégaphone, qu'il peut utiliser

<sup>552</sup> Une version est disponible en ligne sur le Vimeo de Samuel Bianchini, <https://vimeo.com/75058234>, consulté le 30/01/23.

pour parler, crier ou donner des ordres. Plus le volume de la voix sera élevé, plus les golfeurs vont s'agiter. Des voitures de golf vont arriver, déposer d'autres golfeurs, le rythme va s'accélérer. En regardant le détail des mouvements, on voit ces personnages effectuer des mouvements en boucle, dans le vide, l'agitation devient assez vite absurde. Lorsque l'on interrompt le flux amplifié de la voix, les nombreux golfeurs repartent, tout se calme et l'on revient à la situation de départ.

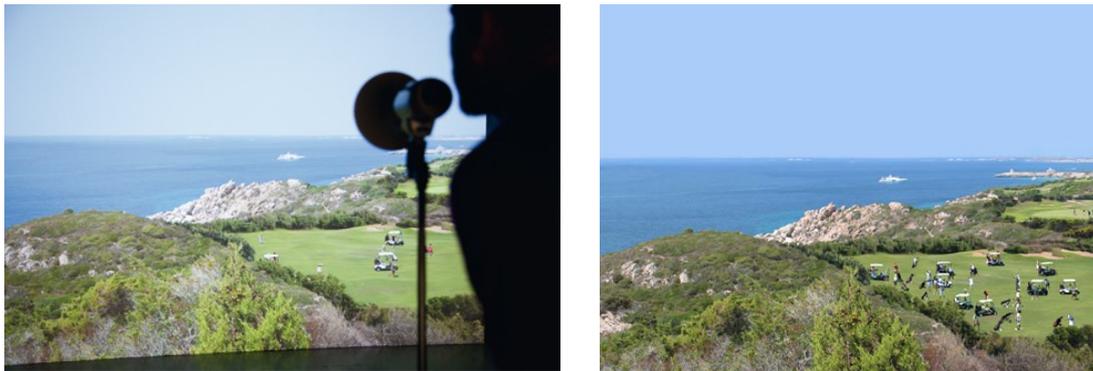


Fig. 87 : Samuel Bianchini, *Temps Libre*, 2004  
© Samuel Bianchini / dispotheque.org

L'image animée est ici esclave (au sens informatique, la relation esclave-maître) de la voix du visiteur. Celui-ci, acteur de la scène, se retrouve dans un rôle se situant à mi-chemin entre un réalisateur de film et un gardien de prison. Pour que l'agitation se mette en place, il faut sans cesse parler, sinon il y a retour. Une addiction se crée entre le spectateur et la vidéo.

### ***Stratifications du réel par la générativité***

Le cadre choisi, avec son arrière-plan beau et apaisant, devient vite le lieu de tensions induites par le fourmillement désynchronisé des personnages-motifs. Techniquement, les golfeurs et les voitures ont été détournés numériquement, puis clonés, c'est-à-dire convertis en objets informatiques dont les instances (ou occurrences) deviennent indépendantes et peuvent chacune se voir affecter un comportement par programmation. En veillant à conserver l'échelle et le relief, il devient ainsi possible de générer des situations dont les développements se basent sur les interactions d'un participant,

ici l'intensité du signal sonore reçu. L'image animée n'est plus un bloc, mais une superposition de microvidéos manipulables, mises en scène sur un arrière-plan<sup>353</sup>.

Pour les images initialement filmées, finalement, la problématique est celle d'une engrammation : comment les analyser (les découper en éléments distincts), les indexer pour pouvoir les remonter/les rejouer ? Et ce processus de démontage devient particulièrement complexe lorsqu'il nécessite une déconstruction de l'image, non plus au photogramme (ou frame) près, mais dans l'image<sup>354</sup>.

Comparé à un jeu vidéo, le trouble opéré entre la saisie du réel et sa « mise en générativité » est extrêmement efficace. Le contraste entre l'humour qui y est instillé et le propos en arrière-plan nous évoquent par ailleurs deux films du cinéaste d'animation Norman McLaren. D'une part *Neighbours* (1952), film réalisé en stop motion avec des prises de vue réelles, et dans lequel deux voisins se disputent de plus en plus violemment au sujet d'une fleur ayant poussé sur la limite de leurs terrains respectifs. D'autre part *Canon* (1964), qui, dans sa dernière partie, met en scène sur fond noir plusieurs occurrences d'un policier canadien dont les mouvements se répondent dans un canon audiovisuel rendu possible par ce même principe d'isolation et de superposition. Le court-métrage *Tango* (1981), de Zbigniew Rybczynski, peut également être cité. Dans celui-ci, un espace clos se remplit peu à peu de personnes, qui entrent et sortent par les portes ou les fenêtres. Tous les personnages sont des détourages (des isolations) qui ont été ensuite superposés dans une même composition. Cela donne la sensation troublante, comme dans *Temps Libre*, d'une forme hybride entre film et animation synthétique. Contrairement à ces exemples d'animation, il n'y a toutefois pas dans *Temps libre*, de résolution, de fin. Cela constitue d'ailleurs une caractéristique de ce que Jean-Louis Boissier, artiste et théoricien de l'interactivité, nomme les « moments interactifs<sup>355</sup> ». On y éprouve une

353 Ce type d'architecture est tout à fait comparable aux « calques » dans les logiciels Photoshop, Illustrator, After Effects, ou bien encore ceux des « pistes » dans un logiciel de son. Simplement, les calques-pistes ne sont pas fixés et exportés pour être rassemblés en un bloc (l'image ou le morceau de musique). Le mixage reste ouvert aux manipulations d'un opérateur, qui peut être l'artiste (dans le cadre d'une performance par exemple), ou un utilisateur, comme dans *Temps Libre*.

354 Samuel Bianchini, « Reverse Writing. De l'écriture audiovisuelle à l'écriture interactive : prises de vues et mises en scène à l'écran », in Jean-Marie Dallet (dir.), *Cinéma, interactivité et société*, Université de Poitiers & CNRS/VDMC, 2013, p. 137.

355 « Les moments interactifs n'ont pas à proprement parler de début et de fin. Ils ont des entrées sans doute [...], mais se perpétuent ensuite dans des boucles sans fin, oscillantes ou circulaires, qui s'in-

sensation mixte, entre le plaisir de contrôler et la déception de réaliser que cela n'a pas d'impact, pas de quelconque finalité.

### ***Variabilité et performance***

La performance désigne, pour Samuel Bianchini, des expériences se trouvant entre situation et action, et dont les objectifs ne sont pas définis au préalable ; leurs résolutions appartiennent au public, à qui est confié un rôle d'opérateur.

L'installation *Temps libre* (2004) est construite selon cette approche et elle revêt un caractère particulier qui permet d'expliquer comment geste et langage se conjuguent dans l'acte de performance. En effet cette installation invite le public à produire des sons pour activer l'image de golfeurs. Ici le geste doit être entendu non pas comme langage dont on tenterait une reconnaissance sémantique, mais comme fait oral dont les sons comportent des dimensions physiques (hauteurs, attaques, volumes, etc.) qui permettent l'interprétation : par exemple — trop brièvement, car cela est plus subtil dans les faits — plus l'intonation (analysée à partir de ces dimensions sonores) est vive et forte, plus on estime que le spectateur donneur d'ordres est décidé, donc plus il agit sur l'image<sup>356</sup>.

Comme dans *The Waves*, le visiteur joue un rôle. Il est dans une relation de co-présence avec l'œuvre, il « performe » avec elle, à la manière des *DJs*. Si la situation d'attente est proche de la projection d'un film cinéma muet, le visiteur, en criant, crée de « l'audiovisualité ». Il devient la voix off d'un film qu'il manipule. Le son, détecté et analysé informatiquement<sup>357</sup>, va créer un rapport de transduction avec l'image : il y a une conversion de paramètres sonores en paramètres d'animation. Cette ouverture du dispositif à une intervention extérieure est fondamentale, car elle offre au son et à l'image de nouveaux moyens de créer des relations.

---

versent, bifurquent, se déversent dans d'autres boucles, selon les actions conjointes de la mise en scène et de la lecture » (Jean-Louis Boissier, « Le moment interactif », in Jean-Louis Boissier, *La relation comme forme. L'interactivité en art, op. cit.*, p. 249).

356 Renée Bourassa, Louise Poissant (dir.), *Avatars, personnages et acteurs virtuels*, Québec, Presses de l'Université du Québec, coll. « Esthétique », 2013, p. 107.

357 Nous parlerons, dans la troisième partie, des commandes vocales qui se fondent sur une reconnaissance de la langue.

## 2.2 Épuisement du fragment audiovisuel : du saut interactif à la microstructure

### 2.2.1 La figure du « saut » dans les arts interactifs

Dans le chapitre précédent, les manipulations de l'axe-temps que sont l'inversion ou le ralenti sont perçues de manière globale par le participant à l'expérience interactive. Elles s'appliquent à un même fragment, qu'il soit saisi en temps réel (la *Time Machine*) ou préparé par l'artiste (*The Waves*). Nous allons à présent étudier comment le numérique autorise le programme à devenir la tête de lecture programmée d'un même fragment audiovisuel, pour y créer des épuisements.

#### *Le saut interactif*

Dans ses *figures de l'interactivité*<sup>358</sup>, Jean-Marie Dallet pose la figure du « saut » comme l'un des fondamentaux de l'interactivité. Il s'agit de donner à un opérateur, par le logiciel, la possibilité d'avancer ou de revenir sur la ligne de temps d'un matériau audiovisuel.

Appuyer sur le bouton de la souris, alors que le curseur est placé sur la vidéo, entraîne l'utilisateur à passer de segment en segment, à « zapper », à sauter d'amorce en amorce. [...] Cette figure de saut est donc proche du « zapping » télévisuel ou plus précisément du « flipping », à la différence près qu'elle ne s'appuie pas sur l'aléatoire, sur la rencontre d'images fortuites mais qu'elle opère sur une substance commune à tous les segments d'images<sup>359</sup>.

Si cette description est focalisée sur l'image, on comprend néanmoins bien que le saut s'exerce sur une substance commune, c'est-à-dire un matériau préparé dans lequel on navigue temporellement. Dans le cas d'un matériau audiovisuel, cela induit que le saut ne brise pas la synchronicité entre l'image et le son qui préexiste au sein du matériau de base. Les deux composantes restent en permanence solidaires l'une de l'autre (à la

358 Jean-Marie Dallet, *La Notion de figure dans les arts interactifs*, thèse de doctorat en arts plastiques, sous la direction de Jean-Louis Boissier, Université Paris 8, Saint-Denis, 2001.

359 Jean-Marie Dallet, glossaire du rapport d'activité du programme *Figures de l'interactivité*, [http://www.dallet.net/PAGEStheorie/figures\\_DAP/figures\\_DAP\\_1/B\\_r\\_DAP\\_7.html](http://www.dallet.net/PAGEStheorie/figures_DAP/figures_DAP_1/B_r_DAP_7.html), consulté le 30/01/23.

manière du ralenti de *The Waves*<sup>360</sup>, qui s'applique aux deux composantes). Ces sauts peuvent être préparés par l'artiste et opérés par une personne, ou bien opérés par l'artiste lui-même, en situation de performance devant un public. D'une certaine manière, ces sauts préparés et/ou exécutés, en temps réel peuvent être considérés comme une forme de montage éphémère.

Dans ce concept de saut, il y a le principe sous-jacent du déplacement instantané d'une tête de lecture. À la fin des années 1800, le bras du gramophone permettait déjà, pour un enregistrement sonore sur disque, de réaliser physiquement cette opération. En revanche, du côté de la pellicule cinématographique d'abord, puis de la bande magnétique dans les années 50, un défilement — aussi mécanisé et rapide qu'il put l'être — était obligatoire pour naviguer dans le matériau sonore, visuel, ou audiovisuel. Ainsi l'accès instantané à un photogramme dans un corpus fixé n'a pu se faire que par le numérique. Notons tout de même que des systèmes proches des jouets optiques du précinéma<sup>361</sup>, mais à base de cylindres similaires à ceux du phonographe, pouvant contenir plus de 40 000 images microscopiques, ont été imaginés par Edison et son assistant Dickinson à l'époque<sup>362</sup>, mais que pour aucun le développement ne semble s'être concrétisé.

Des artistes se sont emparés assez tôt de cette problématique de l'accès. L'une des pièces les plus emblématiques est probablement *Random Access* (1963), de Nam June Paik. Dans cette installation, il dispose et oriente aléatoirement différents morceaux de bande magnétique audio sur un mur, chacun étant séparé des autres pour éviter les superpositions.

---

360 Voir au chapitre précédent, section 2.1.4, avec l'idée de « soudure perceptive ».

361 Le zootrope ou le zoopraxiscope, par exemple, fonctionnait avec des disques sur lesquels étaient disposées en cercle des images dont le défilement rapide donnait l'illusion d'une séquence animée.

362 Lev Manovich, *Le langage des nouveaux médias*, op. cit., p. 136.

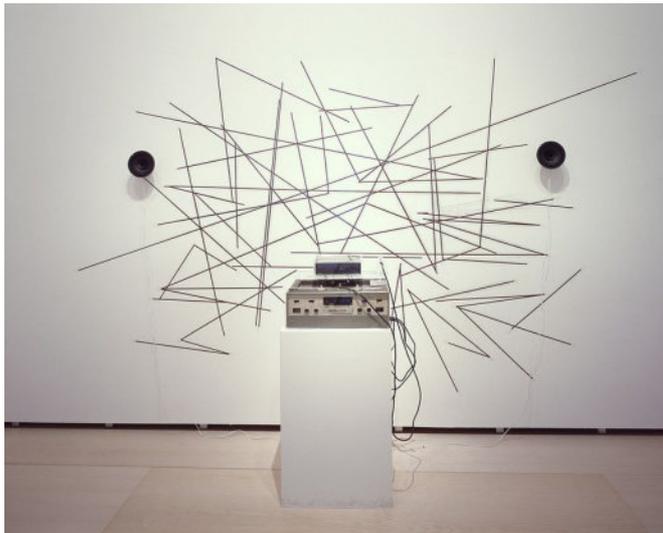


Fig. 88 : Nam June Paik, *Random Access* (1963)  
© Guggenheim Museum

De part et d'autre de la composition murale, des haut-parleurs reliés à un système audio diffusent le son. La lecture du son se fait à l'aide d'une tête de lecture câblée et donc mobile qu'un visiteur passe sur les différentes bandes. Celui-ci compose alors son propre morceau à partir du matériau de base, délinéarisé et dépendant de ses choix de déplacement.

## 2.2.2 Barbara Lattanzi. Un cinéma structurel algorithmique



Fig. 89 : Barbara Lattanzi, *Optical De-dramatization Engine (O.D.E.) applied in 40-hour cycles to Thomas Ince's 'The Invaders', 1912 (2006)*

© Barbara Lattanzi / wildernesspuppets.net

Les sauts interactifs, que nous avons évoqués dans la section précédente, opèrent sur la microstructure du fragment audiovisuel. Ainsi dès lors que le numérique s’immisce dans les dispositifs de cinéma-vidéo, ce qui se passe entre les images — l’Entre-images<sup>363</sup> — devient fondamental, comme le souligne Jean-Louis Boissier :

Cet entre-images est la substance même du cinéma. Le cinéma, dans sa nouvelle configuration numérique, se découvre une série de variabilités par lesquelles l’interactivité s’infiltré, avec lesquelles elle se construit. Ce qui s’offre d’emblée, c’est ce qui se passe entre les images. [...] un premier facteur de variabilité réside donc dans la cadence modulable de l’affichage des photographes, non seulement sur le mode connu de l’accélééré ou du ralenti, mais sur le mode d’un écart pertinent, affectant le rythme des images, la finesse de leurs effets descriptifs et, qui plus est, dans la maîtrise de l’ordre même de l’apparition des images<sup>364</sup>.

Barbara Lattanzi est une artiste dont le travail a débuté dans le champ de l’art internet. Elle explore ces questions de microstructures audiovisuelles et d’entre-images à travers une démarche de software art<sup>365</sup>. La singularité des logiciels-œuvres de Lattanzi — qu’elle caractérise par l’expression *idiomorphic software* — réside non pas dans la multiplicité de possibilités offertes, mais dans la transposition de principes de montage — empruntés ou inventés — en algorithmes, offrant ensuite à l’utilisateur une palette de possibilités performatives. L’œuvre n’est plus le matériau monté en soi, mais le logiciel, à travers les principes qui en sont au cœur. Nous étudions ici deux familles de logiciels conçus par Lattanzi. La première est une série de « logiciels-hommages » au cinéma structurel, créée en 2002. La seconde, à partir de 2006, explore l’étirement temporel de *found footage*, par l’utilisation de l’application *Optical De-dramatization Engine* (O.D.E.).

363 En référence au théoricien Raymond Bellour qui analyse dans ses ouvrages les passages de l’image entre photographie, cinéma et vidéo, Raymond Bellour, *L’Entre-images : photo, cinéma, vidéo*, Paris, La Différence, coll. « Les Essais », 2002.

364 Jean-Louis Boissier, « Propositions pour un cinéma interactif expérimental », in Jean-Marie Dallet (dir.), *Cinéma, interactivité et société*, op. cit., p. 157.

365 Pratique artistique qui place le logiciel et son code comme œuvre, qui considère « la programmation informatique comme la matière même de la pratique artistique » (Andreas Broekmann, « L’esthétique du Software Art », in David-Olivier Lartigaud, *Art++*, op. cit., p. 94-95).

### **Première série de logiciels : traduire le cinéma structurel en algorithmes**

La première famille, de 2002, est une suite de trois logiciels<sup>366</sup> : *HF Critical Mass*, *EG Serene*, *AMG Strain*. Ceux-ci se fondent respectivement sur la transposition de principes structurels<sup>367</sup> de montage utilisés dans les films *Critical Mass* (1971) de Hollis Frampton, *Serene Velocity* (1970) de Ernie Gehr, *Strain Andromeda The* (1992), d'Anne McGuire.



Fig. 90 : Hollis Frampton  
*Critical Mass* (1971)  
© Anthology Film Archives



Fig. 91 : Ernie Gehr  
*Serene Velocity* (1970)  
© MoMA



Fig. 92 : Anne McGuire  
*Strain Andromeda The* (1992)  
© Mubi

Le principe de manipulation est assez similaire. L'utilisateur charge une vidéo dans le logiciel (un simple glisser-déposer) puis une interface, visible à l'écran sous la vidéo, permet en temps réel de manipuler les paramètres principaux. Selon le logiciel : unité de séquençage, sauts interactifs, décalages de synchronisation entre image et son. Chaque nom de logiciel est constitué des initiales du réalisateur et d'un ou de premiers mots du titre.

*HF Critical Mass* (2002)<sup>368</sup> est à la fois une œuvre et un outil destiné à être utilisé dans un contexte d'improvisation audiovisuelle. Il s'agit d'une œuvre logicielle de manipulation numérique développée par Barbara Lattanzi, et dont l'algorithme rend hommage aux procédures de montage utilisées par le cinéaste expérimental Hollis Frampton pour son film *Critical Mass (Hapax Legomena III)*, qui date de 1971 et qui fait partie de la série *Hapax Legomena*. Celui-ci a pour seul matériau de base une dispute

<sup>366</sup> Scott MacKenzie, Janine Marchessault (dir.), *Process Cinema. Handmade Film in the Digital Age*, Montréal, McGill-Queen's University Press, 2019, p. 490.

<sup>367</sup> Dans les années 1950 et 1960, des cinéastes expérimentaux tels que Peter Kubelka, Paul Sharits ou Michael Snow explorent le structuralisme qui « consiste en somme à réduire une œuvre à ses propriétés inhérentes (structures), qui génèrent à leur tour le contenu du film (Paul Young, Paul Duncan [dir.], *Le cinéma expérimental*, Paris, Taschen, 2009, p. 105).

<sup>368</sup> Il s'agit de la version 1, développée sous Director. La version 2 date de 2015 et a été réalisée grâce au logiciel Max/MSP/Jitter.

improvisée et filmée entre une femme et un homme<sup>369</sup>. S'opère dans le film un montage particulier d'image-son qui vient recomposer la dispute, en décalant ou resynchronisant le son et l'image. D'une certaine manière, la structure intrinsèque du film illustre, à travers les relations entre image et son, des formes relationnelles du couple. Frampton a travaillé le montage afin que la chronologie initiale ne soit plus l'axe principal, mais il déplace la narrativité sur un mode plus délié, avec des retours en arrière et des sauts en avant en saccade (comme le fera plus tard Martin Arnold). La première version de *HF Critical Mass*, le logiciel-hommage au film de Frampton, développée par Barbara Lattanzi, permettait de reproduire l'algorithme de sauts, mais ne pouvait pas décaler le son et l'image<sup>370</sup>.

*EG Serene*, deuxième logiciel de la série, se base sur le film *Serene Velocity* (1970) de Ernie Gehr. Le film de Gehr est constitué d'une suite d'images fixes, prises de vues cadrées d'un même couloir. Passées à cadence rapide, mais dans un ordre différent de celui, naturel, qui serait celui de l'avancement (ou du zoom) progressif de la caméra, le spectateur se voit entraîner dans un mitraillage de variations de cadrage qui avance et recule sans arrêt, et qui semble jamais n'arriver. Le logiciel de Barbara Lattanzi, *EG Serene*, remédialise ce principe en proposant à l'utilisateur de jouer avec ces sauts dans le temps.

Pour le troisième logiciel, et par une mise en abyme encore plus profonde, Lattanzi a travaillé sur un film qui est déjà lui-même le réemploi d'un film plus ancien. Dans son film *Strain Andromeda The* (1992), la réalisatrice Anne McGuire a en effet isolé tous les plans-séquences du film de science-fiction *The Andromeda Strain* (1971) de Robert Wise, puis les a remontés en partant du dernier, jusqu'au premier, c'est-à-dire dans l'ordre inverse de la chronologie du film. Il s'agit d'un procédé très similaire à celui de

---

369 Le film présente, dans ses premières minutes (son amorce), un écran noir avec des voix qui se succèdent et s'entremêlent, comme pour annoncer d'emblée que l'on n'assistera pas à une narration classique. Les images animées et les sons vont ensuite se répéter, se casser, et la synchronisation entre image et son qui se manifestait régulièrement va peu à peu se désintégrer, voir donner l'impression de s'inverser, la voix de l'homme doublant la femme, et réciproquement. Les saccades sont basées sur les mots prononcés.

370 Barbara Lattanzi nous a confirmé que ce décalage était trop complexe à obtenir dans la première version. Elle tenait en revanche à ce qu'il soit possible dans la version 2, car cela constituait une caractéristique fondamentale du film d'Hollis Frampton. Un logiciel de Lattanzi plus récent, *Unwriting*, encore plus orienté improvisation, permet, en temps réel, et à partir d'une base de séquences audiovisuelles, de naviguer en temps réel sur l'image et sur le son. Il y est possible de désynchroniser la bande son de la bande image, soit complètement, soit grâce des décalages.

Piotr Kowalski dans sa *Time Machine*<sup>371</sup>, mais appliqué à une unité plus longue, de nature intrinsèquement cinématographique : le plan<sup>372</sup>. Le logiciel de Lattanzi est donc programmé pour offrir la possibilité à un utilisateur de faire glisser un fichier audio-vidéo afin de le faire lire en séquençage inverse. Différents paramètres sont ajustables, mais la programmation consiste en partie à calculer la durée de tronçons audio-vidéo, puis d’emmener la tête de lecture, à rebours, aux différents points de repère<sup>373</sup>. Lattanzi reproduit ainsi dans l’algorithme de son logiciel *AMG Strain* le procédé utilisé par McGuire pour qu’un utilisateur puisse l’appliquer sur n’importe quelle vidéo qu’il charge dans le logiciel.

Ce que les cinéastes structuraux ont exploré de manière artisanale — mettre en lumière la structure du film — est donc traduit par Lattanzi en algorithme (au sens de la description schématique, en langage courant, d’un principe), puis en programme qui s’exécutera dans le contexte d’une interface personnalisée. Pas de base de données complexe, c’est souvent un seul fragment audiovisuel<sup>374</sup> qui se retrouver soumis à une multitude de sauts programmés de la tête de lecture virtuelle pour se retrouver recomposé.

Pour créer ses logiciels, Lattanzi a utilisé Macromedia Director (et le langage Lingo), Macromedia Flash (et son langage ActionScript). Sans rentrer dans le détail, ces deux logiciels multimédia ont été très importants à cette période des années 2000, car ils permettaient, à partir d’un environnement graphique standard constitué d’une scène graphique et de palettes, de créer des narrations interactives en croisant le principe de placement d’éléments sur une *timeline* : formes vectorielles, textes, images, vidéos, sons.

---

371 Voir au début de ce chapitre.

372 Le film *Memento* (2000), de Christopher Nolan, utilise ce type de principe d’inversion chronologique.

373 Ce que l’on nomme communément les *cuepoints* en anglais. Souvent symbolisés par des drapeaux dans les logiciels de montage, ils permettent de marquer des moments importants des flux d’images et de sons.

374 Quand Barbara Lattanzi utilise elle-même son logiciel, dans le contexte d’une performance ou d’une installation, elle choisit fréquemment comme base des vidéos à caractère politique, comme celle en 2015 d’une manifestation de colère à Minneapolis à la suite du décès de Jamar Clark, abattu par la police (Barbara Lattanzi, *Critical Mass applied to Black Lives Matter* (2018), disponible en ligne sur le Vimeo de Barbara Lattanzi, <https://vimeo.com/274382533>, consulté le 30/01/23). Lattanzi propose aussi des séries de performances collectives comme *Do Video Objects Have Tails?* (2002), dans lesquelles des artistes utilisent son logiciel-interface, mais avec leurs propres séquences vidéo.

```

on doCriticalMass me
  member(videoMemNm).sound = 1
  if controllerBarSpr > 0 then
    if tempStartFlag = 1 then
      -- viewer is controlling the playback
      -- the QT video is at the beginning
      -- The following condition signals time to begin the rewind cycle:
      if videoSpr.movieTime > sprite(controllerBarSpr).width then
        tempStartFlag = 0
        if flagMovieRate = 1 then
          checkForwardAmount
        else
          if flagMovieRate = 0 then
            checkReversalAmount
          end if
        end if
      end if
    end if
  end if

  if videoSpr.movieTime >= stopTheAlgorithm - (forwardAmt ) then
    tempStartFlag = 1
  end if
end if

```

Fig. 93 : Barbara Lattanzi, extrait du code source de *HC Critical Mass* © Barbara Lattanzi / *wildernesspuppets.net*

Un principe similaire à ceux des logiciels linéaires de montage, mais ici il était possible d'ajouter des *scripts* (des instructions en code) permettant entre autres potentialités, de programmer des sauts (en ActionScript l'instruction *gotoAndPlay* [20] permettait par exemple d'envoyer la tête de lecture de la *timeline* à l'image 20, à partir d'un clic, d'un survol ou toute autre action possible sur l'interface) en donnant des numéros d'image. Il était donc possible de réaliser une interface personnalisée, avec des boutons ou des éléments graphiques permettant à l'utilisateur d'agir directement sur les médias<sup>375</sup>.

Lattanzi fournit, dans une démarche d'open source, met à disposition son logiciel sur un site, le maintient<sup>376</sup>, et fournit le code. Elle spécifie que même si elle utilise Lingo (Director), le code peut être adapté ou traduit en d'autres langages capables de naviguer dans des médias.

### **Deuxième série de logiciels : étirer le temps**

Le principe d'étirement du temps par sauts réguliers en avant ou en arrière sur la ligne de temps est poussé à son extrême dans un autre logiciel de Lattanzi, *Optical De-dra-*

375 Toute cette génération de logiciels, parce qu'ils pouvaient directement exporter l'interface sous forme d'applications autonomes (les CD-ROM par exemple pour Director) ou de modules intégrables sur une page web (Director et Flash) ont permis de démocratiser la création de logiciels artistiques.

376 C'est-à-dire que comme les développeurs de logiciels traditionnels, elle corrige les bugs et met à jour ses applications régulièrement. Elle peut aussi les reprogrammer — ou les faire reprogrammer —, dans un autre langage. *HF Critical Mass* a par exemple été traduit en Max/MSP, a priori pour des raisons d'obsolescence de Director et de Flash.

*matization Engine* (O.D.E.). Ce système lui permet, toujours à partir d'une seule vidéo ou film, d'étirer artificiellement la durée grâce à deux techniques croisées : d'une part la duplication d'images (toujours par des sauts temporels), d'autre part grâce des zooms et dézooms lents parfois (couplés à un déplacement) dans l'image<sup>377</sup>.

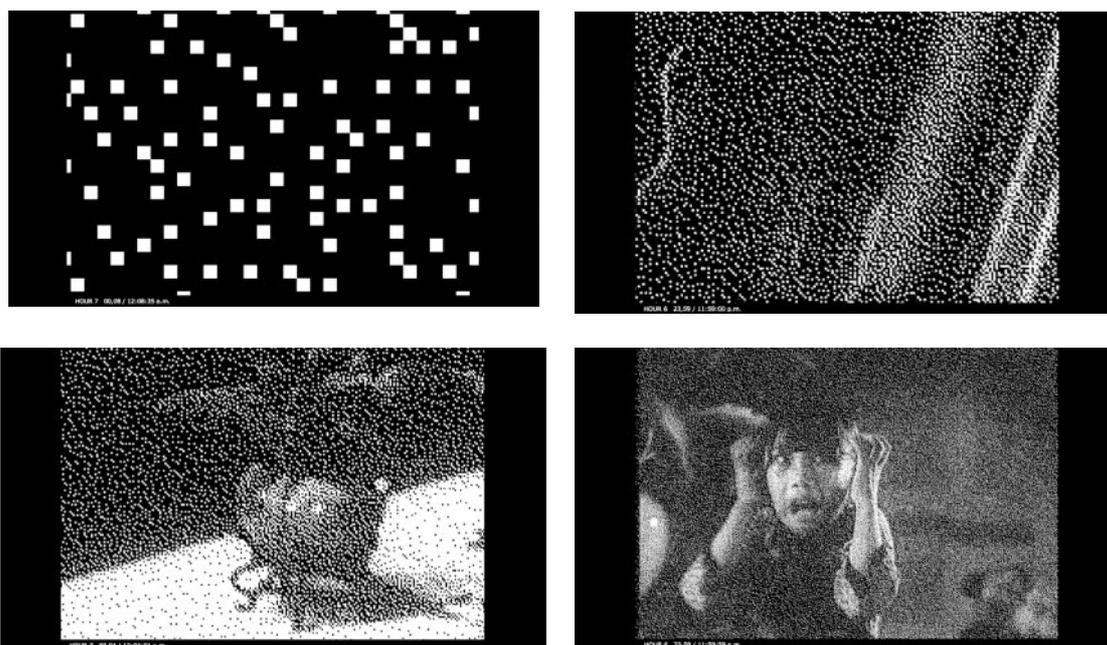


Fig. 94 : Barbara Lattanzi, *Optical De-dramatization Engine (O.D.E.) applied in 15-hour cycles to Ma-Xu Weibang's 'Ye-ban gesheng' (Song at Midnight), 1937* (2008)  
© Barbara Lattanzi / Vimeo

Les images sont par ailleurs pixellisées, de façon à ce qu'un zoom très profond ne montre que des carrés blancs sur fond noir par exemple. Lattanzi a créé deux configurations du logiciel, avec deux films différents. Ces versions étaient disponibles en ligne<sup>378</sup>. Une fois lancées, elles généraient des versions longues et déstructurées (« dé-dramatisées » pour reprendre le terme de Lattanzi) à travers des cycles de plusieurs dizaines d'heures.

La première pièce générative se base sur un film muet de Thomas Ince, *The Invaders* (1912). Son titre, long mais explicite, est *Optical De-dramatization Engine (O.D.E.) ap-*

<sup>377</sup> Ce procédé rappelle les générations automatiques de certains logiciels photo, comme Photos sur Mac OS X, qui peuvent générer automatiquement et lire un diaporama animé en provoquant des déplacements et des zooms sur la photo elle-même pour éviter un statisme. Une musique est d'ailleurs choisie automatiquement dans notre bibliothèque, pour venir également combler le vide du silence.

<sup>378</sup> Elles fonctionnaient sous Flash, sous la condition que le *plugin* correspondant (une extension) soit chargé dans le navigateur.

*plied in 40-hour cycles to Thomas Ince's 'The Invaders', 1912*<sup>379</sup>. La seconde date de 2008 et se nomme *Optical De-dramatization Engine (O.D.E.) applied in 15-hour cycles to Ma-Xu Weibang's 'Yeban gesheng' (Song at Midnight), 1937*<sup>380</sup>. La première pièce est muette (car le film original l'est également) mais dans la seconde, le son est aussi décomposé, comme l'image, en microfragments et recomposé. Par ailleurs, si l'application web est lancée alors que minuit sonne, la chanson phare du film se lance<sup>381</sup>. Ces films génératifs sont parfois installés sous forme de diptyque. Ici l'œuvre est à la fois le logiciel et le résultat de ce qui est créé par l'artiste.

### 2.2.3 Le *scratching* ou l'art de la tête de lecture

Les sauts programmés de Barbara Lattanzi, lorsqu'ils suivent des logiques de retour avant et retour arrière dans la timeline, peuvent s'apparenter à des formes de *scratching* audiovisuel. Ce procédé, chez les *DJs*, consiste à effectuer des sauts et des va-et-vient dans les sillons du vinyle. Ces techniques prennent leurs origines à différents moments. Dans les années 1950 notamment, l'ingénieur Pierre Schaeffer met au jour la technique du sillon fermé<sup>382</sup>, qui préfigure celle du *scratch* et de l'utilisation de la boucle.

#### *L'influence de la musique hip-hop*

Le breakbeat, puis le hip-hop sont nés de cette pratique du *scratching*<sup>383</sup>. Grandmaster Flash innove par exemple avec la technique du « back-spinning » (le retour en arrière rapide d'un passage sans perte de rythme), ou bien en introduisant les premières boîtes à rythmes au sein du dispositif de platines. De par cette maîtrise accrue de l'art de la tête de lecture, il devient facile pour ces artistes de changer très rapidement et

379 Une version est disponible en ligne sur le Vimeo de Barbara Lattanzi, <https://vimeo.com/idiomorphics>, consulté le 30/01/23.

380 Une version est disponible en ligne sur le Vimeo de Barbara Lattanzi, <https://vimeo.com/141938927>, consulté le 30/01/23.

381 Une capture de cette situation est disponible sur le compte Vimeo de Barbara Lattanzi.

382 À l'époque, quand un disque 78 tours était abîmé, le « saphir sautait et répétait sans cesse le même fragment [...]. Schaeffer pensa alors que ce qui était un défaut d'enregistrement ailleurs pouvait ici être considéré comme un avantage : la possibilité de mettre les sons en boucle ». (Peter Manning, « La musique concrète et ses appareils », in Martin Kaltenecker, Karine Le Bail [dir.], *Pierre Schaeffer. Les constructions impatientes*, Paris, CNRS Éditions, 2012, p. 142-146).

383 Parmi ces musiciens, citons les pionniers DJ Herc, Grandmaster Flash ou Afrika Bambaataa.

sans coupure brusque leurs disques, ces derniers ne jouant plus alors qu'un rôle d'échantillon dans un ensemble devenu composition.

L'idée de base du breakbeat était aussi simple que révolutionnaire. Jusqu'alors, la plupart des DJ, si l'on excepte les pionniers du disco, avaient considéré un morceau comme une unité et l'avaient encensé comme tel. Herc et ses épi-gones considéraient les chansons comme des sortes de carrières auxquelles ils pouvaient arracher des pierres pour la construction de leurs propres œuvres. [...] En ayant simplement eu l'audace de revendiquer l'histoire de la musique comme matériau de sa propre création, le DJ s'éleva au statut d'auteur, c'est-à-dire d'initiateur, créateur et fondateur d'un phénomène nouveau lui appartenant en propre<sup>384</sup>.

Au milieu des années 1980, la pratique du *scratch* gagne le champ de la vidéo. Des groupes de musique industrielle tels que Cabaret Voltaire ou Throbbing Gristle intègrent de l'image animée dans leurs performances scéniques, en général des montages de *found footage* de vidéos violentes ou politiques. Parmi eux, les Duvet Brothers, puis au début des années 1990, les groupes Emergency Broadcast Network ou Coldcut, qui vont développer la pratique du scratch vidéo en live. Ces nouvelles façons de manipuler l'image et du son vont conduire au développement d'un grand nombre d'outils dédiés.



Fig. 95 : Laurie Anderson et son Viophonograph  
© Olivier Lussac / artperformance.over-blog.fr

---

384 Ulf Poschardt, *DJ Culture*, Paris, Éditions Kargo/L'Éclat, 2002, p. 172.

Parmi d'autres exemples, l'artiste Laurie Anderson produit à la fin des années 1970 une expérimentation de scratch avant l'heure, avec un violon revisité.

En 1978, elle crée le Viophonograph en intégrant un diamant sur son archet et, sur le corps du violon, un disque qui tourne et qui comporte une succession de phrases musicales qu'elle joue à sa guise en déplaçant le diamant. « C'est un des premiers instruments avec lequel j'ai pu faire du scratch », dit-elle<sup>385</sup>...

Parmi de nombreuses expérimentations avec son violon dont elle « étend » constamment les possibilités, elle conçoit en 1984, avec le pionnier de la synthèse informatique Max Mathews, une interface capable de naviguer dans les sons d'un disque dur.

### ***L'instrument comme outil de navigation***

Cette facilité accrue de sauter et de naviguer dans des images est également devenue centrale dans les versions contemporaines de la performance *Violin Power* de l'artiste Steina. Dans la proposition initiale (dont les versions s'étalent entre 1969 et 1978) présentée sous forme de vidéo linéaire de démonstration<sup>386</sup>, Steina, musicienne de formation, applique grâce à son violon différents effets de distorsion sur sa propre image diffusée en direct sur un moniteur, jusqu'à ce que l'effet prenne le pas sur la figuration. On ne voit plus alors que la représentation graphique en mouvement d'un signal audio. Le son du violon est capté par un microphone, puis passé à travers un générateur d'effets vidéo appelé le *Rutt/Etra Scan Processor*.

---

<sup>385</sup> Jacqueline Caux, « Les violons de Laurie Anderson », in *Art Press* n° 285, 2002.

<sup>386</sup> Une version est disponible en ligne sur le Vimeo de The Kitchen, <https://vimeo.com/555723409>, consulté le 30/01/23.



Fig. 96 : Steina, *Violin Power* (1969-78)  
© vasulkakitchen.org

Au début des années 1990, en explorant les possibilités numériques, elle fait évoluer sa performance. Présentée comme une « performance interactive », le son du violon est toujours audible, mais il dissimule un système non plus dédié à distordre l'image, mais destiné à naviguer dans des corpus de vidéos stockées sur un support numérique<sup>387</sup>.

Alors, quand on a découvert l'appareil LaserDisc qui, outre l'accès instantané à toute image fixe sur le disque, pouvait exécuter des fonctions comme avancer ou reculer rapidement, sauter, arrêter sur une image, faire une pause, grâce au contrôle MIDI, alors tout ce qu'il me fallait, c'était un violon MIDI. J'avais déjà tenté de faire bouger une bande dans un magnétoscope avec un violon acoustique et c'était possible. Mais sans l'accès aléatoire, je ne pouvais faire bouger la bande que dans un secteur limité. Maintenant que le LaserDisc est désuet, j'ai déplacé ces fonctions d'exécution de la performance dans l'ordinateur sous forme de films QuickTime grâce au logiciel Image/ine. Aujourd'hui, tout ce qu'il me faut, c'est un portable et mon violon MIDI<sup>388</sup>.

Sur ce violon électrique à cinq cordes appelé ZETA, certains points d'appui sur les cordes déterminent des accès à des séquences d'image, tandis que d'autres règlent la vitesse et la direction du défilement des images. L'une des cordes permet d'assigner

<sup>387</sup> Nous dépassons ici légèrement le cadre théorique du chapitre, car si Steina peut effectuer des sauts au sein d'un même fragment, elle peut également naviguer dans un corpus d'éléments hétérogènes. Nous évoquerons ces questions dans le troisième chapitre, dédié aux bases de données.

<sup>388</sup> Yvonne Spielmann, « Entretien avec Steina », Fondation Langlois, <https://www.fondation-langlois.org/html/f/page.php?NumPage=416>, consulté le 30/01/23.

aux autres différentes fonctions. Steina, en 1997, est ensuite passée à un autre système basé sur le logiciel Open Source *Image/Ine*, outil destiné à la performance live<sup>389</sup>.



Fig. 97 : Steina, *Violin Power The Performance* (1992)  
© vasulka.org

Il apparaît également possible pour l'artiste de passer le violon en sourdine afin de laisser le son de la vidéo (lui-même distordu ou ralenti suivant ce qui est appliqué à la vidéo).

### ***Défilement programmé. La remédiation du va-et-vient structurel***



Fig. 98 : Martin Arnold, *Passage à l'acte* (1993)  
© Crac Occitanie

---

389 « Steina, *Violin Power*, 1978 », Fondation Langlois, <https://www.fondation-langlois.org/html/f/page.php?NumPage=485>, consulté le 30/01/23.

Nous pouvons également retrouver cette pratique du défilement contrôlé dans les films expérimentaux de Martin Arnold. Sa trilogie constituée de *Pièce touchée* (1989), *Passage à l'acte* (1993)<sup>390</sup> et *Alone* (1998), nous donne à voir et à entendre des pièces filmiques constituées de *remix* temporel d'un fragment ou d'une scène unique issue de films hollywoodiens. Si le son dans *Pièce touchée* ne consiste qu'en quelques boucles relativement simples et ajoutées après le montage visuel final, à partir de *Passage à l'acte* les recompositions temporelles fonctionnent par appariage image/son. C'est-à-dire que les mêmes manipulations temporelles sont effectuées simultanément et systématiquement sur des portions d'image-son ; la synchronisation image-son originale est entretenue, c'est l'ensemble qui est manipulé.

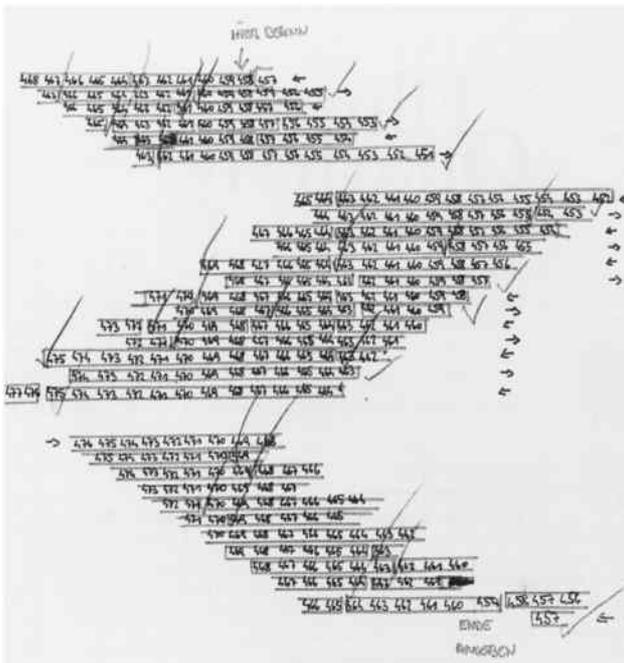


Fig. 99 : Partition de *Pièce touchée*

© Martin Arnold / Researchgate Max Neupert

La partition complète fait près de 200 pages. Elle nous montre le principe de va-et-vient dans le déroulement du film. Un principe similaire est appliqué dans *Passage à l'acte* et *Alone*, mais sur l'image et le son synchrones.

Il conserve également, dans une certaine mesure, la chronologie de la scène ; le début du film est bien le début de la scène, la fin celle de la scène. Dans l'intervalle s'opèrent

<sup>390</sup> Une version est disponible en ligne sur le Vimeo de L'Éclat, <https://vimeo.com/398345542>, consulté le 30/01/23.

des retours, des avances, des boucles, des saccades qui créent de nouveaux rythmes et étirent la durée du film. Sont mis en potentialité un nombre infini de combinaisons possibles, de sauts dans le temps, de relations audiovisuelles.

Pour *Alone* (1997-1998), Martin Arnold a utilisé l'ordinateur pour l'aider à mettre au point ses structures en défilement avant et arrière. Ceci n'est pas anodin, car cela lui a permis d'une part de voir de manière instantanée les résultats, d'autre part d'obtenir une plus grande précision du montage sonore, notamment en défilement inverse<sup>391</sup>.

L'étirement du temps induit par les va-et-vient que Martin Arnold a patiemment fabriqués, à la fin des années 1980, par un travail manuel sur la pellicule devient avec le numérique le quotidien des monteurs. À l'aide d'une molette associée à leur logiciel de montage, ils naviguent en temps réel dans des séquences linéaires, par sauts temporels et va-et-vient. Ils regardent, reviennent, passent, écoutent, écoutent à nouveau.

#### 2.2.4 *L'interprète*, d'Eric Vernhes. Vers le sampling audio-vidéo

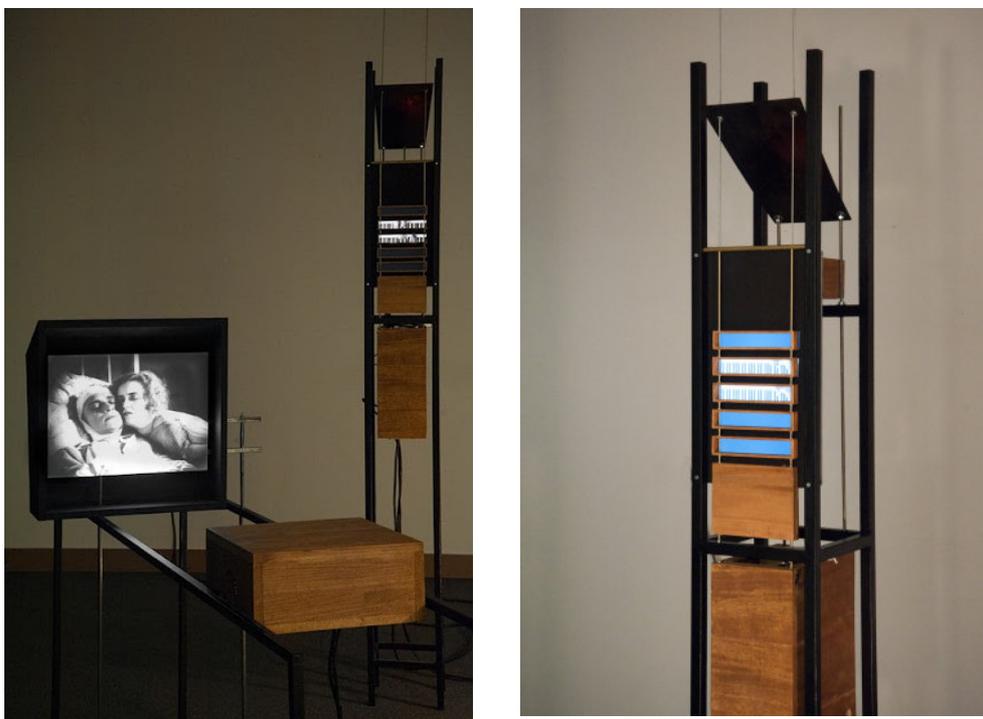


Fig. 100 : Eric Vernhes, *L'interprète* (2010)

<sup>391</sup> Extrait d'une interview de Martin Arnold. « Le cinéma et le corps : un art du luxe », in Yann Beauvais, Jean-Michel Bouhours (dir.), *Monter-sampler : l'échantillonnage généralisé*, Paris, Centre Pompidou/Scratch Projection, 2000, p. 79-81.

Comme de nombreux procédés utilisés par les DJ et les VJ, la technique du scratching audiovisuel, dès lors qu'elle est utilisée trop systématiquement, devient un « effet ». Dans la pièce que nous allons étudier ici — *L'interprète* (2010)<sup>392</sup> d'Eric Vernhes —, cela vient au contraire servir un propos original, entremêlant une histoire dramatique et les pratiques musicales de l'artiste.

Les installations de l'artiste contemporain Eric Vernhes font coexister l'image et le son de manière singulière. Il s'agit souvent d'ensembles plastiques, sortes de dispositifs-sculptures qui intègrent dans leur architecture un ou plusieurs petits écrans, ainsi que des systèmes de diffusion sonore discrets. Le cinétique et le numérique y dialoguent : pour exemple dans *Machine à écrire* (2016) se côtoient et communiquent — grâce à des interfaces et des programmes numériques — un mécanisme d'horlogerie (pour un pendule), un piano mécanique, et des images génératives rappelant le film abstrait des années 1920, *Symphonie Diagonale* de Viking Eggeling<sup>393</sup>. Cette hétérogénéité se comprend aussi dans les descriptions des œuvres, qui mentionnent les matériaux utilisés ; par exemple « Acier, bois, ordinateur, écran ».

En ce qui concerne *L'Interprète*, elle est tout autant œuvre de réemploi que création originale : l'artiste y remixe deux films anciens, en noir et blanc, tout en apportant ses propres séquences audiovisuelles. Les pièces cinématographiques préexistantes, *Orlacs Hände* (1924) de Robert Wiene et *Mad love* (1935) de Karl Freund sont deux adaptations cinématographiques du même roman : *Les mains d'Orlac* (1921), de Maurice Renard. Celui-ci raconte l'histoire d'un pianiste (Stephen Orlac) qui, suite à un accident, se fait greffer les mains d'un criminel venant d'être exécuté. Autour d'une histoire d'amour et de différents événements, les mains essaient de prendre le pouvoir sur le héros. Les séquences audiovisuelles tournées par Eric Vernhes le présentent, quant à elles, en train d'improviser lui-même au piano. Le cadrage sur le clavier, en vue de dessus, ne montre, du corps de l'artiste, que les mains.

---

392 Une captation est disponible en ligne sur le site d'Eric Vernhes, <https://www.ericvernhes.com/project/linterprete/>, consulté le 30/01/23.

393 Voir au début de la première partie, autour de la musique visuelle.



Fig. 101 : Eric Vernhes, *L'interprète* (2010)  
© Eric Vernhes / ericvernhes.com

Ces deux natures de vidéo sont aussi celles qui définissent l'espace de l'installation audio-visuelle. Deux parties principales la composent. La première (à gauche sur la première image au tout début de cette section) est un écran sur lequel les extraits entremêlés des deux films de *found footage* sont projetés et soumis à des manipulations temporelles. L'autre (à droite sur la même image) est une colonne, située en retrait. Conçue « comme un instrument de musique à part entière<sup>394</sup> », elle présente six mini-vidéos placées les unes au-dessus des autres. Les petits écrans (il s'agit en réalité d'un seul écran LED, par-dessus lequel sont fixés une série de cadres en bois), que l'on pourrait appeler les pistes, présentent les différents moments d'improvisation d'Eric Vernhes sur son piano. Celles-ci ne sont pas toujours actives, parfois deux s'allument en même temps, parfois trois, parfois toutes.

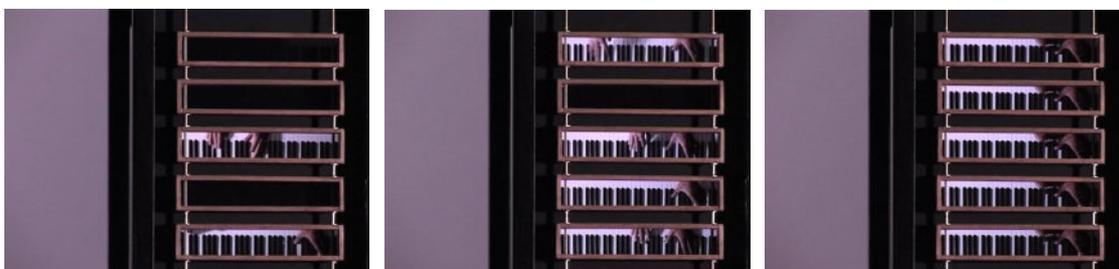


Fig. 102 : Eric Vernhes, *L'interprète* (2010)

<sup>394</sup> Texte de présentation de la pièce par l'artiste, <https://www.ericvernhes.com/project/linterprete/>, consulté le 30/01/23.

Capture vidéo de différents moments d'improvisation, diffusés sur la colonne.  
© Vimeo / Eric Vernhes

Cette barre de vidéos de claviers de piano est en quelque sorte le contrôleur du montage des films. Pendant que les vidéos de piano sautent, s'inversent (Vernhes a utilisé un logiciel de *Vjing* pour créer le montage multipiste), la projection des films à gauche est soumise aux mêmes transformations. Il y a transfert — transduction — des sauts temporels imposés aux improvisations vers les échantillons cinématographiques. Au début le film remonté est en extrême ralenti (des effets ont été utilisés) et l'une des pistes de piano se joue à vitesse normale. Puis tout s'emballe. Les deux entités vidéo ont beau être séparées dans l'espace, un couplage structurel s'établit.



Fig. 103 : Eric Vernhes, *L'interprète* (2010)  
© Vimeo / Eric Vernhes

Le son en revanche (celui des morceaux de piano manipulés) n'est diffusé que par la colonne qui présente les microvidéos. Pour Vernhes, cette colonne symbolise la situation du pianiste, qui accompagnait les films au temps du cinéma muet. Si les appareils de projection sont placés à l'intérieur d'élégants caissons en bois, la diffusion sonore est également travaillée. La partie inférieure de la colonne masque un caisson de basse, et un plan incliné réfléchit les aigus<sup>395</sup>. Il y a donc séparation spatiale des images et du son, mais une partie des opérations de montage et de manipulation effectuées

395 *Ibid.*

par le programme sur les pistes de piano se répercutent sur le montage du film, créant une forme audiovisuelle assez addictive, assez proche des films de Martin Arnold. Le piano étant aussi bien présent dans les films que dans les pistes vidéo d'improvisation, le trouble est d'autant plus fort.

Eric Vernhes travaille la matière audiovisuelle en prenant soin d'accorder la même importance à l'image qu'au son. L'une des dynamiques sous-jacentes des pièces de l'artiste est d'ailleurs une égalité entre les éléments sonores et visuels.

Michel Chion disait qu'un son est une chose, une image une autre et qu'ensemble ils forment une entité à part, l'audio-vision. Quand l'ordinateur a été capable de gérer les deux, j'y ai vu l'occasion de les travailler simultanément. Je n'ai jamais cessé depuis<sup>396</sup>.

À ce titre, mais aussi de par le sujet du roman qu'elle met en scène, la pièce *L'interprète* peut être considérée comme une greffe d'audio-vision : un réemploi de films pré-existants et de séquences originales d'image et de son. Le numérique vient « brancher » ces composantes hétérogènes pour les faire fonctionner ensemble.

### **2.3 Esthétique des bases de données d'images et de sons**

Les procédés de manipulation temporelle — inversion, ralenti, accéléré ou saut — évoqués dans les deux chapitres précédents relèvent de la manipulation intensive de quelques fragments audiovisuels. Lorsque l'artiste s'engage dans la création de formes exigeant une multiplicité de séquences distinctes, se pose alors la question fondamentale, explorée notamment par Lev Manovich, de la base de données<sup>397</sup>. Le saut interactif, qui permettait de se positionner à différents moments d'un segment particulier, se voit confronté à des corpus complexes d'images et de sons. Le programme n'est alors plus seulement tête de lecture. Il va pouvoir, à la volée, analyser, classifier et agréger des médias, mais aussi agencer en direct la macrostructure d'une forme artistique.

---

396 Samantha Deman, « La mécanique sensible d'Eric Vernhes », ArtsHebdoMédias, 2016, <https://www.artshebdomedias.com/article/mecanique-sensible-deric-vernhes/>, consulté le 30/01/23.

397 Lev Manovich, *Le langage des nouveaux médias*, op. cit.

### 2.3.1 La base de données dans l'art

Pour les chercheurs Bertrand Gervais et Jean-Marie Dallet, les bases de données sont « les nouveaux palais de mémoire<sup>398</sup> ». Elles enregistrent, retiennent et préservent les données. Elles se constituent en véritables architectures, semblables à des palais des glaces. Elles cherchent à s'effacer et à nous perdre dès lors qu'elles se complexifient, pour devenir des boîtes noires au fonctionnement opaque. Il est donc important d'en explorer les rouages.

Or, faire parler les données, c'est bien l'un des enjeux des architectures de mémoire. Qu'est-ce qui a été retenu et enregistré, de quelle manière l'intégrité des données a-t-elle été préservée, comment organiser ces données pour ne pas les perdre dans le lot, et ensuite, chercher à déterminer ce qu'elles signifient ? Car les données ne signifient pas toutes seules, il faut les interroger<sup>399</sup>.

Dans sa thèse<sup>400</sup>, l'artiste Andreas Kratky fait remonter l'origine du terme « base de données » à 1964. Celui-ci est alors utilisé par un groupe de techniciens opérant dans le champ militaire. Il y avait là l'idée d'un système centralisé de données, accessible à plusieurs personnes en même temps. Pour le chercheur et artiste Lev Manovich, la base de données est une forme culturelle, nouvelle forme symbolique de l'ère informatique<sup>401</sup>, qui met en œuvre une symbiose entre données et algorithmes. Il la définit comme une sorte de collection structurée d'éléments qui donnent lieu à des « objets néo-médiatiques », dont les éléments « ne s'organisent pas en séquences thématiques, formelles ou autres, mais constituent plutôt des ensembles d'informations qui ont toutes la même importance<sup>402</sup> ». L'enjeu est alors d'organiser convenablement ces éléments de façon à ce qu'ils puissent être récupérés rapidement et efficacement par un programme informatique<sup>403</sup>. Le cinéaste expérimental et théoricien Malcolm Le Grice<sup>404</sup>,

---

398 Jean-Marie Dallet et Bertrand Gervais, « Architectures et palais de mémoire », in Jean-Marie Dallet, Bertrand Gervais, (dir.), *Architectures de mémoire*, Dijon, Les presses du réel, 2019, p. 10-11.

399 *Ibid.*, p. 12-13.

400 Andreas Kratky, *Les auras numériques : pour une poétique de la base de données*, thèse en arts plastiques et sciences de l'art, sous la direction d'Anne-Marie Duguet, Paris, Université Paris 1 Panthéon-Sorbonne, 2013, p. 46. Notons que l'auteur de cette recherche a participé à l'élaboration du projet *Soft Cinema* de Lev Manovich, dont nous parlerons dans une section ultérieure de ce chapitre.

401 Lev Manovich, *op. cit.*, p. 394.

402 *Ibid.*, p. 393.

403 *Ibid.*, p. 400.

404 Nous avons parlé de sa pièce *Arbitrary Logic* dans la première partie, section 1.3.2.

dans ses textes consacrés à l'apparition du numérique, a également très tôt pointé cette nouvelle manière de stocker et de récupérer quasi instantanément les données, par assignation d'adresses dans une mémoire-matrice : « le temps nécessaire pour aller d'une adresse à une autre est électroniquement égal quel que soit le numéro d'adresse choisi, tous les emplacements sont conceptuellement équidistants<sup>405</sup> ». Il relève en outre que ce système de mémoire vive, appelé aussi mémoire de travail ou RAM (Random Access Memory), caractéristique du stockage et de l'accès informatique, ouvre des potentialités de montage non linéaire<sup>406</sup>.

Pour Lev Manovich, cette non-linéarité et ce poids égal des éléments qui en constitue la matière définissent la base de données en tant qu'entité qui « représente le monde comme une liste d'éléments qu'elle refuse d'ordonner<sup>407</sup> ». Cet emploi du mot « liste » fait écho très directement à la théorie de la raison computationnelle développée par le philosophe des techniques Bruno Bachimont. Celui-ci cherche, dans son travail, à délimiter les contours de ce que serait une logique propre au numérique. S'appuyant sur le concept de raison graphique développé par l'anthropologue Jack Goody<sup>408</sup>, Bachimont postule que nous sommes passés, avec le numérique, à une raison computationnelle. Selon Goody, lors du passage de la parole à l'écriture, des façons inédites d'envisager la langue sont apparues. Les trois structures conceptuelles que sont la liste, le tableau et la formule en sont les exemples les plus probants : ils sont difficilement transmissibles par l'intermédiaire du langage parlé et constituent donc une raison graphique. Se référant à ce modèle, Bachimont décrit une raison computationnelle qui met en jeu trois nouvelles composantes : le programme, qui se substitue à la liste. Le réseau qui se substitue au tableau. La couche qui se substitue à la formule. Pris au sens de réseau, la base de données peut ainsi s'affirmer comme l'un des nouveaux paradigmes importants du numérique.

La raison graphique a produit la raison classificatoire, la raison computationnelle produit la pensée en réseau et le temps de la prévision. Pour une raison graphique, le réseau n'est pas une structure de l'intelligible : le réseau, échappé

---

405 Malcolm Le Grice, « Cinéma numérique et film expérimental », in Yann Beauvais (dir.), *Le temps des images*, op. cit., p. 216.

406 *Ibid.*, p. 216.

407 Lev Manovich, op. cit., p. 403.

408 Jack Goody, *La Raison graphique, La domestication de la pensée sauvage*, Paris, Minuit, 1977.

pant à la synopsis spatiale de fait de sa complexité, est un labyrinthe où l'on se perd. C'est une figure de l'irrationnel, et non une manière de penser le monde. L'interaction et la communication selon la structure des réseaux sont devenues intelligibles car le calcul permet de réduire la complexité et de parcourir l'ensemble des possibles induit par les réseaux par les programmes qui en spécifient le comportement<sup>409</sup>.

Bertrand Gervais et Jean-Marie Dallet, dans leur circonscription conceptuelle des bases de données — les « architectures de mémoire » —, s'appuient sur cette théorie de la raison computationnelle, c'est-à-dire une « rationalité particulière reposant sur le calcul plutôt que sur l'écriture<sup>410</sup> ». Le chercheur en design Anthony Masure, analysant également la pensée de Bachimont, la prolonge en évoquant la persistance d'une tabularité<sup>411</sup>. Héritée du régime graphique, elle viendrait se matérialiser numériquement dans une logique de réseau, en s'appuyant sur des principes tels que les hypertextes<sup>412</sup>. Ces hypermédiats<sup>413</sup> fonctionneraient alors sur des principes de non-linéarité et de « pensée associative<sup>414</sup> ». Selon N. Katherine Hayles, chercheuse en humanités numériques, le modèle dominant est d'ailleurs celui de la « base de données relationnelle », fondé sur des systèmes de tables, interconnectées par des relations (des liens) entre des champs<sup>415</sup>, porteurs d'associations potentielles. Le chercheur et poète Jean-Pierre Balpe relève également les nombreux points communs existants entre les bases de données et les hyperdocuments, notamment lorsque leur modèle est

409 Bruno Bachimont, « Signes formels et computation numérique : entre intuition et formalisme. Critique de la raison computationnelle », 2004,

[http://www.utc.fr/~bachimon/Publications\\_attachments/Bachimont.pdf](http://www.utc.fr/~bachimon/Publications_attachments/Bachimont.pdf), consulté le 30/01/23.

410 Jean-Marie Dallet et Bertrand Gervais, « Architectures et palais de mémoire », in Jean-Marie Dallet, Bertrand Gervais, (dir.), *op. cit.*, p. 14-15.

411 C'est-à-dire « [...] la possibilité pour le lecteur d'accéder à des données visuelles dans l'ordre qu'il choisit, en cernant d'emblée les sections qui l'intéressent, tout comme dans la lecture d'un tableau l'œil se pose sur n'importe quelle partie, dans un ordre décidé par le sujet » (Christian Vandendorpe, « Linéarité et tabularité », in Christian Vandendorpe, *Du papyrus à l'hypertexte. Essai sur les mutations du texte et de la lecture*, Paris, La découverte, coll. « Sciences et société », 1999, <https://www-cairn-info.ezpaarse.univ-paris1.fr/du-papyrus-a-l-hypertexte--9782707131355-page-41.htm>, consulté le 30/01/23).

412 Anthony Masure, *Design et humanités numériques*, Paris, B42, coll. « Esthétique des données », 2017, p. 123.

413 Ted Nelson invente en 1965 le terme « hypertexte » pour désigner des écrits non linéaires. Les hypermédiats en sont l'extension à des utilisations multimédia (image, vidéo, son).

414 Anthony Masure, *op. cit.*, p. 124.

415 « Les relations entre les éléments de données figurant dans les tables sont construites selon des opérations ensemblistes telles que “insérer”, “effacer”, “sélectionner” et surtout “joindre”, la commande qui permet de combiner des données de différentes tables. Les éléments communs permettent d'établir des corrélations entre les tables. » (N. Katherine Hayles, *Lire et penser en milieux numériques. Attention, récits, technogenèse*, Grenoble, Ellug, 2016, p. 283).

organisé en réseau<sup>416</sup>, de même que Jean-Marie Dallet, qui écrit que « l’hypertexte et les hypermédias présentent une méthode d’organisation de l’information qui procède non plus de façon linéaire, mais par association d’idées en agencant un “réseau vivant” de liens entre les données<sup>417</sup> ».

Pour que la base de données existe, il faut organiser ses éléments. Il est nécessaire pour cela de les catégoriser, de les découper en éléments significatifs. L’artiste et théoricien Samuel Bianchini a pointé ce rôle fondamental de l’indexation.

Sans aller jusqu’à être une forme de notation, au sens musical du terme, l’indexation s’annonce comme une sorte de précomposition. Les fragments forment une matière première, rendue intelligible et organisable grâce à son mode d’indexation. Leur organisation préalable permettra de trouver une correspondance juste avec le code informatique qui les gouverne. Une fois réuni et organisé, un ensemble considérable de données devient base de données : des données mises à disposition de programmes qui pourront y faire appel pour les traiter, les agencer ou les monter<sup>418</sup>.

Mais cela induit toujours, comme le rappelle Hayles, une implication idéologique<sup>419</sup>. Les vocabulaires et taxinomies d’indexation sont l’un des leviers sur lesquels les artistes peuvent s’appuyer pour détourner et modeler les formes. Ils sont intimement liés au fonctionnement du moteur de recherche qui sera mis en œuvre pour récupérer les données ou les médias.

Une autre strate, modelable par les artistes, est celle de l’interface. Elle permet à l’artiste, ou à un utilisateur, de naviguer dans ces corpus. Bon nombre d’artistes, notamment dans le Net art<sup>420</sup>, ont détourné les logiques de la recherche et de l’indexation en

416 Jean-Pierre Balpe, *Hyperdocuments, hypertextes, hypermédias*, Paris, Eyrolles, 1990, p. 14.

417 Christophe Domino, Jean-Marie Dallet, « Une archive active pour l’Institut pour l’abolition de la guerre. Principes pour une collaboration », in Patrick Nardin, Catherine Perret, Soko Phay, Anna Seiderer (dir.), *Archives au présent*, Vincennes, Presses universitaires de Vincennes, coll. « Esthétiques hors cadre », 2017, <https://www-cairn-info.ezpaarse.univ-paris1.fr/archives-au-present--9782842927431-page-195.htm#no10>, consulté le 30/01/23.

418 Samuel Bianchini, « Une opération partagée. Le montage confronté aux technologies de l’interactivité. », in Yann Beauvais, Jean-Michel Bouhours (dir.), *op. cit.*, p. 108.

419 N. Katherine Hayles, *op. cit.*, p. 286.

420 Depuis les début du Web au début des années 1990, le terme désigne les « créations interactives conçues par, pour et avec le réseau Internet » (Jean-Paul Fourmentaux, « Net art », in *Communications* n° 88, 2011, <https://www-cairn-info.ezpaarse.univ-paris1.fr/revue-communications-2011-1-page-113.htm>, consulté le 30/01/23).

inventant des dispositifs singuliers d'agrégation de données, mais, encore une fois, peu ont travaillé sur le côté sonore<sup>421</sup>.

La base de données (locale, ou connectée par un réseau, comme celui d'Internet), couplée au programme (qui permet d'établir des chemins et des configurations possibles), et à l'interface (qui modèle la navigation et la manipulation), est à la base de ce que l'on a pu appeler les hypermédias. Comme Jean-Marie Dallet le préconise, les stratégies d'indexation et d'agrégation opérées par les bases de données doivent être mises en lumière par les artistes et les chercheurs :

Si la base de données est coextensive à la société numérique dans laquelle nous baignons, nous sommes alors tentés de considérer les algorithmes de recherche et leurs scénarios comme des optiques à même de nous permettre des *vedute* multiples, ubiquitaires et mouvantes sur le monde virtuel façonné par les architectures de mémoire. Celles-ci apparaissent aujourd'hui comme les monuments inédits d'une fabrique des regards et des consciences individuelles qu'il nous faut observer et critiquer<sup>422</sup>.

L'irruption des bases de données dans les dispositifs artistiques induit ainsi de nouvelles possibilités de relation entre l'image et le son. L'instantanéité de l'accès offert par le montage non linéaire et les potentialités associatives apportées par l'indexation (manuelle ou automatisée par analyse informatique) constituent deux axes esthétiques centraux au sein des œuvres que nous allons étudier dans les prochaines sections.

---

421 Parmi ces des rares exemples, le moteur *Netomat*<sup>TM</sup> (1999), créé par l'artiste Maciej Wisniewski. Plutôt que d'afficher les résultats d'une recherche sous forme de liste, il décide d'inventer une présentation beaucoup plus singulière. L'utilisateur saisit classiquement les mots-clés d'une recherche, dans une zone dédiée discrète, en bas à gauche de l'écran. Immédiatement l'interface donne à voir et à entendre un flux animé de titres, d'images et de sons relatifs à la recherche. Alors que la plupart des moteurs de recherche comportent des onglets qui séparent et affichent les résultats par nature (articles, photos, vidéos...), *Netomat*<sup>TM</sup> entremêle les médias. Ceux-ci passent de droite à gauche de l'écran, portés ensemble comme par un vent virtuel. Dans ce mouvement infini, aucun clic n'est autorisé, seul le mouvement de la souris permet d'explorer ce flux. L'artiste crée ainsi, très tôt, une forme de langage audio-visuel novateur fondé sur l'interrogation en temps réel d'une base de données (ici le Web), avec l'utilisateur en disc-jockey. Il existait même une version « salon » du moteur, *Netomatheque* (2001), qui permettait à l'utilisateur de naviguer par commande vocale. Un site-hommage existe à cette adresse : <https://www.netom.at/>. Son moteur de recherche n'est pas fonctionnel, mais il permet de comprendre le principe d'interface de navigation. Parmi d'autres, Rachel Greene évoque ce travail dans un ouvrage consacré au « net art » : Rachel Greene, *L'art Internet*, Paris, Thames & Hudson, coll. « L'univers de l'art », 2005, p. 132.

422 SLIDERS\_Lab (Frédéric Curien et Jean-Marie Dallet), « Plan de vol », in Jean-Marie Dallet, Bertrand Gervais, (dir.), *op. cit.*, p. 120-121.

### 2.3.2 Une figure fondamentale. Le décrochage interactif et acousmatique dans *The Erl King*



Fig. 104 : Grahame Weinbren et Roberta Friedman, *The Erl King* (1982-85)  
© Vimeo / Grahame Weinbren

L'avènement de l'ordinateur a donné lieu, dans les années 1980 à de nouvelles formes artistiques qui reposent sur des bases de données d'images et de sons saisis<sup>423</sup>. L'utilisateur peut naviguer, grâce à des interfaces interactives, dans des corpus audiovisuels. Jean-Louis Boissier, artiste et théoricien des médias, rappelle toutefois qu'à la naissance de ces médias interactifs, les enregistrements audio et vidéo étaient encore stockés sous une forme analogique. Le programme informatique, seul, conférait à l'œuvre interactive, par hybridation, sa nature numérique. C'était notamment le cas avec le premier support interactif ayant été utilisé par les artistes, le vidéodisque<sup>424</sup>.

423 Au sens d'enregistrement du réel.

424 Aussi nommé LaserDisc ou laser disc. Ancêtre du CD-ROM et du DVD interactif. Le lecteur LaserDisc avait été conçu, au début des années 1980, comme un équipement grand public, peu coûteux, utilisable avec une simple télécommande. Mais il pouvait aussi être connecté à un ordinateur : « Cette capacité du système au repérage trouve son utilisation la plus intéressante lorsque le lecteur de vidéodisques est relié à un micro-ordinateur qui gère, dans le cadre de programmes plus ou moins complexes, l'appel d'image et le temps de visualisation, et qui peut enrichir l'image vidéo de textes et de graphismes en alternance ou en incrustation » (Juliette Raabe, « Vidéodisque et édition », in Jacques Anis, Jean-Louis Lebrave [dir.], revue Linx, n° 17, « Le texte et l'ordinateur », 1987, [https://www.persee.fr/issue/linx\\_0246-8743\\_1987\\_num\\_17\\_1](https://www.persee.fr/issue/linx_0246-8743_1987_num_17_1), consulté le 30/01/23).

Dans ce beau disque de trente centimètres à l'aspect chromé, se rangent comme autant de sillons circulaires concentriques, à lire au rayon laser, 54 000 images vidéos. Car, à cette étape du devenir numérique de l'image, la vidéo, autrement dit la photographie et le cinéma, reste analogique. L'informatique est là pour trouver, agencer, distribuer ce qui provient d'une saisie. Un vidéodisque se pilote dans un langage qui est ouvert au lecteur, à commencer par sa couche supérieure qui n'est autre que le langage naturel<sup>425</sup>.

L'ordinateur, vis-à-vis des vidéos saisies et stockées, joue donc, aux origines des médias interactifs, le rôle d'un opérateur capable de naviguer quasi instantanément dans un corpus fixé. Parmi ces premières expériences interactives<sup>426</sup>, les œuvres questionnant explicitement la relation entre image et son sont très rares. La bande sonore est pensée comme un accompagnement musical au service d'une narrativité de l'image, ou bien il s'agit du son ambiant de la séquence. Une exception, toutefois : *The Erl King* (1982-85)<sup>427</sup>, de Grahame Weinbren et Roberta Friedman. Avant cette œuvre, les deux artistes avaient déjà réalisé plusieurs films expérimentaux ensemble.

### ***Une interaction tactile***

*The Erl King* est une pièce de vidéo interactive qui s'appuie sur le rapprochement de deux textes : un poème de Johann Wolfgang von Goethe, *Le Roi des Aulnes*<sup>428</sup> (1782), et « Le rêve de l'enfant qui brûle », tiré de *L'interprétation du rêve* (1889) de Sigmund Freud. L'installation comporte deux parties principales : un poste « de commande » (dans lequel un visiteur est assis face à un moniteur) et un espace de projection (un moniteur, placé en hauteur, qui réplique ce qui est affiché sur le moniteur de commande). Nous y voyons et entendons des scènes vidéo, tournées par les artistes et adaptées des deux textes. Le « participant<sup>429</sup> » (auditeur-spectateur-agissant), ou

425 Jean-Louis Boissier, « L'interactif en mots », in Jean-Paul Fourmentraux (dir.), *Images interactives. Art contemporain, recherche et création numérique*, op. cit., p. 15.

426 Parmi elles *Le Bus* (1985), installation vidéo-interactive de Jean-Louis Boissier avec des étudiants de l'Université de Paris 8. Nous parlerons de *Lorna* (1984) de Lynn Hershman Leeson, et de *Portrait n° 1* (1990) de Luc Courchesne, qui utilisent également un système de LaserDisc, dans la troisième partie.

427 Une documentation vidéo est disponible en ligne sur le Vimeo de Grahame Weinbren, <https://vimeo.com/100314200>, consulté le 30/01/23.

428 Dans le récit de Goethe, un enfant accompagné de son père, tous deux à cheval, croit distinguer dans la forêt le Erlkönig (roi des Aulnes). Il tente à plusieurs reprises d'attirer l'enfant. Son père ne le croit pas, et l'enfant finit par mourir, « emporté » par la créature malfaisante.

429 Plutôt qu'utilisateur ou spectateur, Gilles Rouffineau rappelle que Bill Viola, tout comme Frank Popper, avait utilisé le terme « participant » pour définir le visiteur qui à la fois voit, entend, écoute, interagit. Le terme « opérateur » est aussi défini par Rouffineau comme celui « qui met en pratique des pro-

opérateur, peut contrôler le flux narratif directement grâce au moniteur, qui est tactile. Cette notion de participation est notamment décrite par Grahame Weinbren en référence au travail du compositeur John Cage. Celui-ci cherche en effet à mieux répartir, dans son travail, l'expérience musicale entre le compositeur, l'interprète et l'écouter<sup>430</sup>, qui collaborent ensemble. Pour Cage, il n'y a pas d'auditeur passif.



Fig. 105 : Grahame Weinbren et Roberta Friedman, *The Erl King* (1982-85)

© Alain Depocas, Jon Ippolito, Caitlin Jones (dir.), *L'approche des médias variables : la permanence par le changement*, New York / Québec, Guggenheim Museum/Fondation Daniel Langlois, 2003

À gauche, vue globale de l'installation originale. Le moniteur placé en hauteur n'est pas tactile, il duplique simplement ce qui est vu par le participant, pour le reste du public puisse le voir.

À droite, L'espace d'interaction du participant, avec un moniteur tactile.

Quand on touche l'écran pour la première fois, une vidéo démarre. Il s'agit toujours de la même séquence : la captation vidéo d'une chanteuse (la soprano Elizabeth Arnold) qui interprète en allemand le Lied<sup>431</sup> *Erlkönig* (The Erl King) (1888) de Schubert. Si l'on touche à nouveau l'écran, la voix chantée subsiste, mais l'image synchronisée de la chanteuse laisse place à une séquence vidéo autonome, qui illustre l'histoire. Par exemple, un jeune garçon semble fuir quelque chose, et monte en courant les marches

cessus — des opérations — qui peuvent être décrits et sont susceptibles d'être formalisés, transmis et appris » (Gilles Rouffineau, *Éditions off-line : projet critique de publications numériques 1989-2001*, Paris, B42, 2018, p. 117).

430 Grahame Weinbren, « Time Artists: John Cage, Hollis Frampton, Pat O'Neill », 2000, [http://grahameweinbren.com/GW\\_Papers/ICCEssay.html](http://grahameweinbren.com/GW_Papers/ICCEssay.html), consulté le 30/01/23.

431 Un lied, dans la culture musicale allemande, est l'interprétation vocale d'un texte, accompagné par un petit ensemble musical, voire souvent un piano seul, comme ici, avec *Erlkönig*.

d'un escalier. Selon l'endroit de l'écran touché, le scénario diffère. Jean-Louis Boissier en décrit l'usage ainsi :

Le simple toucher de l'écran induit des sauts dans les registres complexes d'images fixes, séquences, textes, musiques et voix. Nul récit n'est visé. Seuls les déclenchements, les superpositions et les ruptures mobilisent le spectateur sur un mode musical et par un sentiment partagé entre le ludique et le désir de structure<sup>432</sup>.

Un autre aspect, qui concerne à la fois l'image et le son, est la capacité de l'œuvre à tenir compte de ce qui a déjà été exploré : « Par exemple, l'utilisation de conditionnels globaux, qui définissent une sorte de carte des données, indique les parties auxquelles on a déjà accédé<sup>433</sup>. Le logiciel ferme certains passages et en rend d'autres accessibles en fonction de cette carte<sup>434</sup> ». Il s'agit d'une logique très proche de celle du jeu vidéo, mais dans un contexte de vidéo interactive.

### ***Le décrochage de l'image et du son***

Chaque nouvelle séquence peut aussi posséder sa propre bande-son, qui vient alors se juxtaposer avec celle de la vidéo musicale du départ. Au fil du scénario, ce sont de nouveaux segments de films, mais aussi de segments de son, qui peuvent surgir. Tout est construit pour que se conserve une continuité cinématographique. Cette potentialité de pouvoir désynchroniser les images et les sons est décrite ainsi par les artistes :

D'autres questions émergent des structures de données, qui dépendent de la séparation de la vidéo et de l'audio. La plus importante d'entre elles est le « décrochage interactif » : une personne touche à l'écran, une nouvelle image

432 Jean-Louis Boissier, *La relation comme forme. L'interactivité en art, op. cit.*, p. 113.

433 Ajoutons que régulièrement, un rectangle blanc filaire s'affiche. Si on le touche, la partie visuelle sera remplacée par la photographie d'une voiture délabrée (qui semble avoir été laissée à la casse). Celle-ci est couverte de tags (produits par les artistes) qui retranscrivent l'une des strophes du poème – par exemple au moment de son chant). Toutes les parties du texte ont été découpées pour produire différentes photographies.

434 Il y a également plusieurs niveaux d'association, directement liés à la narration ou relatifs à des détails présents dans certaines séquences, ce qui crée des imbrications. Le poème de Goethe mentionne par exemple des fleurs sauvages qui fleurissent au bord d'une rivière. Les artistes ont filmé des séquences montrant des tulipes en soie dans les eaux boueuses de L'East River. Autre exemple, la captation de la chanteuse est tournée dans un studio avec des palmiers. On retrouve ces palmiers sur un tee-shirt d'Hawaï dans une séquence (Grahame Weinbren, « The Erl King and its Digital Emulation », 2011, [http://grahameweinbren.com/GW\\_Papers/Emulating%20the%20ErlKing.html](http://grahameweinbren.com/GW_Papers/Emulating%20the%20ErlKing.html), consulté le 30/01/23).

vidéo s'affiche tandis que l'audio continue, et l'écran revient par la suite à la vidéo d'origine. Cette syntaxe est utilisée tout au long de l'œuvre et nécessite beaucoup de paramètres complexes qui sont définis dans le système-auteur et auxquels le système d'exécution donne suite<sup>435</sup>.

Cette opération est possible techniquement grâce à une lecture parallèle de trois LaserDiscs, ce qui permet à tout moment de *switcher* l'image en laissant courir le son. Ce procédé de décrochage est très semblable à celui d'« acousmatisation », proposé par Michel Chion. Fréquemment utilisée dans le montage cinématographique, cette figure audio-visuelle est décrite par le théoricien comme un procédé dramatique visant « à transporter le spectateur-auditeur, à un moment crucial de l'action, dans un lieu extérieur ou éloigné, ne nous laissant plus que le son — devenu acousmatique<sup>436</sup> — pour imaginer ce qui se passe<sup>437</sup> », c'est-à-dire une « opération consistant à nous faire entendre sans voir, après nous avoir permis d'entendre et voir en même temps<sup>438</sup> ». Dans les « chapitres » de *The Erl King*, certaines vidéos servent de référence, de guide. Elles amorcent une scène audio-visuelle, mais leur composante visuelle est susceptible d'être remplacée (par l'action du participant) par une autre, tandis que le son persiste. Il y a décrochage de l'image vis-à-vis de la bande sonore. Elles jouent, en quelque sorte, un rôle de séquenceur de scènes, par l'intermédiaire du son. D'autres vidéos pourraient, quant à elles, être qualifiées d'interchangeables<sup>439</sup>.

435 Jeff Rothenberg, « Grahame Weinbren et Roberta Friedman, The Erl King », In Alain Depocas, Jon Ippolito, Caitlin Jones (dir.), *L'approche des médias variables : la permanence par le changement*, New York / Québec, Guggenheim Museum/Fondation Daniel Langlois, 2003, p. 100-107, version numérique, [https://variablemedia.net/pdf/Erl\\_king.pdf](https://variablemedia.net/pdf/Erl_king.pdf), consulté le 30/01/23.

436 Michel Chion définit la situation acousmatique définit comme celle qui consiste à « entendre sans voir », comme c'est le cas avec le téléphone, la radio, mais aussi avec les bruits de la ville, du voisinage. Ce sont des bruits dont on n'identifie pas visuellement la source. Pierre Schaeffer utilisait ce terme pour définir une musique (la musique concrète) dont l'écoute se portait sur le son lui-même (l'objet sonore), et non plus sa cause. Par ce qualificatif, il faisait référence aux acousmatiques, élèves de Pythagore qui écoutaient le philosophe parler, derrière un rideau (Michel Chion, *Le son, op. cit.*, p. 200-201).

437 Michel Chion, « Glossaire », entrée « Acousmatisation », <http://www.lampe-tempete.fr/ChionGlossaire.html>, consulté le 30/01/23.

438 *Ibid.*

439 Certaines formes interactives du vidéo-clip ont utilisé un principe similaire. La bande-son est fixée (le clip est toujours celui d'un morceau musical linéaire), mais l'utilisateur peut, pendant l'écoute, naviguer, ou interagir avec la partie visuelle.



Fig. 106 : Grahame Weinbren et Roberta Friedman, *The Erl King* (1982-85)  
© Vimeo / Grahame Weinbren

À un segment sonore peuvent ainsi être associées plusieurs images. La décision de changer<sup>440</sup> cette image revient au participant, plutôt qu'au réalisateur-monteur. Une partie du processus de montage<sup>441</sup> est ainsi laissée à l'opérateur. Même si les associations sont préparées, le montage du film est non linéaire.

La place du son est alors primordiale. Tout en créant une continuité narrative, il offre à l'image des potentialités d'interchangeabilité. Le théoricien Gilles Rouffineau, dans son étude sur l'interactivité off-line (par le CD-ROM notamment), avance que les interruptions et les reprises sonores, intempestives ou répétées, produisent des effets mécaniques et artificiels. Un son plus longtemps continu, en revanche, aide à fabriquer la narration, car il joue le rôle de liant<sup>442</sup> entre les fragments<sup>443</sup>.

440 Nous utilisons ce terme plutôt que « choisir », puisque les images possibles ne sont pas connues à l'avance.

441 Le cinéaste Jean-Luc Godard a décrit ainsi son processus de montage entre image et son : « Assis à la table de montage, je commence par visionner les images sans le son. Puis je fais passer le son sans les images. C'est seulement ensuite que je mets les deux ensemble comme ils ont été tournés. J'ai parfois le sentiment que quelque chose ne va pas dans une scène — peut-être irait-elle avec un autre son. Alors je remplace un dialogue par des aboiements, par exemple. Ou j'essaie avec une sonate » (Jean-Luc Godard, cité par Alexandre Castant, in Alexandre Castant, *Planètes sonores. Radiophonie, arts, cinéma*, Blou, Monografik, 2010, p. 190.

442 « Ces objets, qui comportent des fragments temporels, en particulier sonores, construits pour de courtes durées limitent ou refusent à ce moment toute interruption. Il n'y a là rien de paradoxal. C'est au contraire la condition pour que les coupures, raccords et montages imprévus et fréquents, liés à la réactivité des hyperliens de ces programmes, garantissent un temps continu permettant la possibilité de construire un sens. Sur le fond de continuités discursives, même brèves et elles-mêmes fragmentées, les disruptions ont une chance de s'intégrer dans un tout cohérent. Le son a cette propriété de liant particulièrement fort. » (Gilles Rouffineau, *op. cit.*, p. 114)

443 Une illustration radicale de cet effet pourrait être le film *Wavelength* (1967), de Michael Snow qui consiste en un long zoom de 45 min, soutenu par un son sinusoïdal continu dont la fréquence varie du grave à l'aigu. Comme l'explique le réalisateur dans un échange avec l'artiste Christian Marclay, « *Wavelength* est constitué en réalité de plans successifs, et construit de façon très irrégulière. Cela a été fait manuellement. Ce qui sonne l'impression d'un zoom mécanique continu, c'est l'homogénéité du son,

### ***Les narrations parallèles dans une autre pièce interactive : Sonata***

Le principe de décrochage entre l'image et le son est, quelques années plus tard, exploré à nouveau par Grahame Weinbren dans la pièce de vidéo interactive *Sonata* (1993)<sup>444</sup>. Celle-ci est en partie inspirée par une nouvelle de Léon Tolstoï, *La Sonate à Kreutzer* (1889).



Fig. 107 : Grahame Weinbren et Roberta Friedman, *Sonata* (1993)

© Vimeo / Grahame Weinbren

Le principe de navigation réside toujours dans l'interaction avec l'image<sup>445</sup>, mais avec un degré de liberté supplémentaire. L'un des cas les plus singuliers est celui du traitement en diptyque « élastique » de deux scènes au déroulement parallèle. Dans la première séquence, à gauche, une femme joue du piano accompagnée d'un homme au violon. Dans la seconde, à droite, un autre homme s'agite, seul, en gros plan, dans une autre pièce. Il s'agit du mari qui soupçonne le violoniste d'être l'amant de sa femme. Tandis qu'il entend continûment la musique jouée, le participant peut étendre la surface de l'une des vidéos qui vient passer par-dessus l'autre (et inversement). Il s'agit en quelque sorte d'un split-screen adaptatif dont on pourrait faire bouger la ligne de séparation verticale de gauche à droite ou de droite à gauche. Un « rideau » interactif qui permet au participant de mettre en regard les deux scènes. Enfin, si celui-ci pointe

---

parce que le son, lui est continu » (« Conversation entre Michael Snow et Christian Marclay », in Jean-Pierre Criqui [dir.], *Christian Marclay : Replay*, catalogue d'exposition, Paris, Cité de la musique/Réunion des musées nationaux, 2007, p. 131).

444 Une capture vidéo est disponible en ligne sur le Vimeo de Grahame Weinbren, <https://vimeo.com/265094385>, consulté le 30/01/23.

445 Dans les versions les plus récentes, le dispositif est très élaboré, il s'agit d'un cadre à travers lequel le visiteur voit l'image projetée. Un système de capteurs à faisceaux infrarouges permet de détecter la position du doigt.

le doigt dans le coin supérieur gauche de l'écran, le programme fait directement sauter l'action jusqu'au moment du meurtre de la femme par son mari devenu fou.

Le système de mon équipement me permet de traiter le son et l'image comme des éléments séparés, de sorte que le son puisse changer pendant que la vidéo continue, et (le plus souvent) inversement. Le son (musique ou dialogue) continue souvent pendant que l'image change en raison de l'entrée du spectateur, donnant une sensation de fluidité à l'expérience et de continuité au récit [Nous traduisons]<sup>446</sup>.

La pièce est très complexe — nous ne faisons qu'en effleurer les fonctionnements — mais que ce soit dans celle-ci ou dans *The Erl King*, la question d'un décrochage interactif (que les artistes nomment aussi « coupure interactive ») qui préserve une continuité narrative, donnée par le son ou l'image, est essentielle.

### 2.3.3 *Soft Cinema*. Le cinéma algorithmique de Lev Manovich

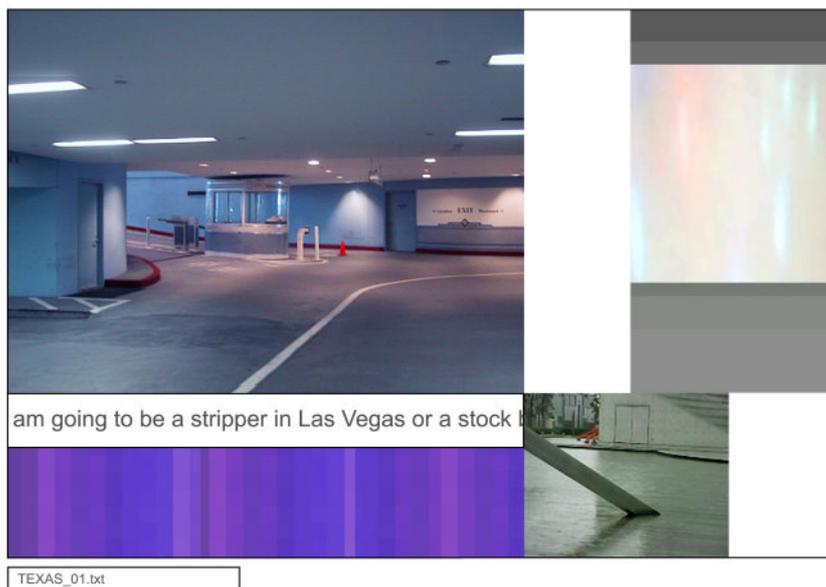


Fig. 108 : Lev Manovich, « Texas », *Soft Cinema* (2002-2005)  
© digitalartarchive.at

446 « My system of equipment allows me to treat sound and image as separate elements, so that sound can change while video continues, or (more commonly) vice versa. The sound (music or dialogue) often continues while the image changes because of viewer input, giving a sense of fluidity to the experience and continuity to the narrative » (Grahame Weinbren, « Random Access Rules », in Kay Hoffmann, Thomas Elsaesser [dir.], *Cinema Futures: Cain, Abel Or Cable? The Screen Arts in the Digital Age*, Amsterdam, Amsterdam University Press, 1996, [https://www.academia.edu/644375/Random\\_Access\\_Rules](https://www.academia.edu/644375/Random_Access_Rules), consulté le 30/01/23).

Dans la section précédente, nous montrions combien *The Erl King* était un exemple pionnier de l'utilisation artistique d'une base de données : pour construire sa narration audio-visuelle interactive, la pièce autorise des décrochages de l'image et du son, en donnant à l'utilisateur l'opportunité de substituer des images à d'autres, parmi celles de la base.

Le théoricien des média Lev Manovich, qui a analysé ces rapports entre narrativité et base de données dans les objets néomédiatiques<sup>447</sup>, les a aussi explorés dans le cadre de recherches artistiques. En 2002, il conçoit au ZKM, pour l'exposition *Future Cinema*<sup>448</sup> un dispositif de cinéma algorithmique, nommé *Soft Cinema (2002-2005)*<sup>449</sup>. L'installation principale<sup>450</sup> est une chambre de visionnage qui abrite un dispositif audiovisuel classique (un moniteur et des haut-parleurs). L'écran (divisé en plusieurs zones) nous donne à voir des séquences vidéos, tandis qu'une voix raconte un récit et que se jouent des accompagnements musicaux. Tout en partageant des principes avec *The Erl King*, *Soft Cinema* pousse plus loin les logiques médiatiques propres à la base de données (Manovich les nomme *Database Narrative*).

### ***Une composition générative***

À chaque nouvelle lecture, le programme découpe l'écran selon des règles préétablies, démarre un segment de fiction sonore (on entend la voix qui lit le texte), charge dans certaines zones des séquences vidéos (certains aplats, colorés ou blancs, sont laissés vides), et joue une musique piochée dans la base de données. Parfois, des segments de texte provenant de la fiction viennent se placer dans des zones dédiées, à la manière des bandeaux des chaînes d'information. Le titre de la séquence est indiqué en bas à gauche, dans une zone dédiée. Une composition multimédia spatialisée, constituée de

---

447 Voir la section 2.3.1 de ce chapitre consacré aux bases de données.

448 *Future Cinema. The Cinematic Imaginary after Film*, ZKM, Karlsruhe, 2002-2003. Commissaires : Jeffrey Shaw et Peter Weibel.

449 Une capture vidéo est disponible en ligne sur le Vimeo du V2, <https://vimeo.com/44602928>, consulté le 30/01/23.

450 Le projet est un ensemble d'expérimentations dans différents domaines. Par exemple, plusieurs espaces de projection ont été imaginés par des designers. Nous nous concentrons toutefois ici aux différentes versions de ce que le visiteur peut voir et entendre dans l'espace de diffusion.

voix, de textes, d'images et de musique s'autofabrique ainsi sous les yeux de l'utilisateur, qui n'a aucune possibilité d'agir.

Concernant la nature des médias utilisés, l'artiste s'est appuyé sur trois corpus. Le premier, qui préexiste à *Soft Cinema*, est un ensemble de fictions écrites par Manovich depuis 1998 (le projet est nommé *GUI* comme *Global User Interface*). Ce sont des nouvelles qui se déroulent chacune dans des lieux différents (Berlin, Riga, Tokyo...), que l'artiste n'a jamais visités. Pour *Soft Cinema*, ces fictions ont fait l'objet d'une lecture enregistrée, puis ont été découpées en segments narratifs qui guident des chapitres. Le deuxième corpus est un ensemble de vidéos tournées par l'auteur, dans les endroits décrits dans les fictions. Pour qu'une cohérence formelle émerge de ces tournages, Manovich s'est fixé un certain nombre de règles, s'inspirant de celles du Dogme 95<sup>451</sup> : une caméra portable, des plans-séquences, des lieux non spectaculaires, pas d'effet. Les sujets filmés sont généralement des endroits sans présence humaine directe, comme des zones industrielles, des voitures circulant sur des routes... Par ailleurs, certaines séries de séquences sont tournées à partir d'un point de départ (l'intersection de deux rues par exemple) et suivant des cercles concentriques. À un même endroit peuvent coexister une dizaine de vidéos<sup>452</sup>. Plusieurs cercles peuvent ainsi se croiser, des séquences peuvent appartenir à deux cercles, créant la potentialité d'une bifurcation entre deux lieux. Le troisième corpus, enfin, est un ensemble de musiques composées pour l'œuvre par des artistes, par exemple un corpus de samples de George Lewis, les morceaux du CD *The Quick and The Dead* (1999) de DJ Spooky, musicien et producteur de hip-hop et Scanner<sup>453</sup>, ou bien des musiques de Jóhann Jóhannsson. Tous sont des musiciens-compositeurs dont le travail s'inscrit aux frontières des musiques contemporaines, avant-gardistes et expérimentales.

---

451 En 1995, les réalisateurs Lars von Trier et Thomas Vinterberg ont rédigé un manifeste s'opposant à l'esthétisation des films. Ils préconisaient un certain nombre de règles à adopter pour la réalisation, par exemple : « Le son doit être produit en même temps que les images, et inversement », « La caméra doit être portée à la main. Tout mouvement, ou non-mouvement, possible avec la main est autorisé », « Le film doit être en couleurs. Un éclairage spécial n'est pas acceptable. », « Tout traitement optique ou filtre est interdit », etc. (Le Monde diplomatique, « Manière de voir » n° 151, février-mars 2017, <https://www.monde-diplomatique.fr/mav/151/VINTERBERG/57082>, consulté le 30/01/23).

452 D'autres vidéos, moins nombreuses, sont des plans fixes animés, des images de synthèse 3D, ou des diagrammes qui viennent enrichir la composition.

453 Robin Rimbaud, dit Scanner, est connu pour avoir à ses débuts utilisé des conversations téléphoniques détournées dans ses compositions électroniques expérimentales.



Fig. 109 : Lev Manovich, *Soft Cinema* (2002-2005)  
© digitalartarchive.at

### ***Un système de règles et d'associations***

Chaque nouvelle lecture offrira une toute nouvelle version de l'œuvre, mais il ne s'agit pas pour autant d'une lecture aléatoire, une option qui n'intéresse pas Manovich : « Si il [l'utilisateur] se contente d'accéder à différents éléments l'un après l'autre et au hasard, comme c'est généralement le cas, il n'y a aucune raison de penser que ces éléments constituent le moindre récit<sup>454</sup> ». Les logiques sont donc à chercher dans le programme<sup>455</sup>, et surtout dans les modes de classification des différents corpus de médias. Comme dans *The Erl King*, quelques vidéos sont systématiquement synchronisées à la narration, tandis que d'autres sont en quelque sorte « interchangeables ». Manovich pousse néanmoins le système de la base de données plus loin, car il classifie ses séquences selon différents critères : la localisation, le type de mouvement, la lumière. Une sorte de taxinomie similaire à celles des tags sur les réseaux sociaux. De cette manière, le programme peut agréger des séquences qui ont des liens entre elles, mais selon des critères plus ou moins perceptibles pour l'utilisateur. Les vidéos jouées à un moment donné dépendent toujours des précédentes. Concernant la relation entre image et son, une part importante se joue au niveau de l'indexation. En effet, le corpus de samples de George Lewis est classifié selon les mêmes critères que les séquences vi-

454 Lev Manovich, *op. cit.*, p. 407.

455 Le système est programmé en Lingo, le langage qui était lié au logiciel Macromédia Director, utilisé à l'époque pour créer des applications multimédia interactives telles que le CD-ROM.

déo, par exemple : « vue de la ville », « espace avec un écran », « intérieur privé », « intérieur public », « objet »<sup>456</sup> ... Il y a donc des corrélations d'image-son sous-jacentes.

Concernant le découpage visuel génératif — qu'il nomme macro-cinéma —, Manovich évoque les compositions géométriques de Mondrian, mais aussi les habillages télévisuels des chaînes d'information ou de finance, découpées en zones qui offrent différentes natures de médias : textes défilant, diagrammes, images, vidéos... Les ensembles de moniteurs agrégeant des vidéos de surveillance, mais aussi les films utilisant la technique du split-screen<sup>457</sup> sont des références pour ce projet. La spatialisation visuelle, qui était absente de *The Erl King* (la pièce fonctionnait par séquençage), et s'annonçait subtilement dans *Sonata* (la division de l'écran en diptyque interactif) est ici explorée comme une nouvelle variabilité : celle de la coexistence du temps et de l'espace<sup>458</sup>.



Fig. 110 : Lev Manovich, *Soft Cinema* (2003)  
Détail d'une visualisation de la base de données,  
qui présente des prises de vue géolocalisées à Berlin  
Les couleurs sont liées à différentes natures de vidéo  
© softcinema.net [archive.org]

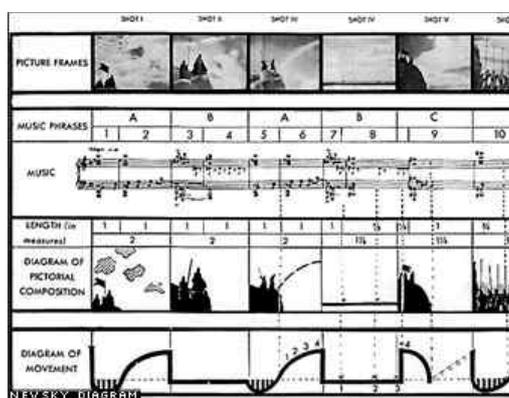


Fig. 111 : Sergei Eisenstein, structure de montage d'une  
séquence de son film *Alexander Nevsky* (1939)  
© softcinema.net [archive.org]

456 Lev Manovich, *Soft Cinema. Navigating The Database*, livret du DVD, Cambridge, MIT Press, 2005.

457 Parmi les exemples emblématiques au cinéma, citons *Timecode* (2001) de Mike Figgis, ou les films de Peter Greenaway comme *The Pillow Book* (1996). Dans les hypermédias, la pièce web interactive *My Boyfriend Came Back from the War* (1996) de l'artiste Olia Lialina, fondée sur des *iFrame* (des zones de l'écran autonomes), met en scène une narration non linéaire sur un mode d'écran divisé (voir par exemple la description de la *Net Art Anthology*, <https://sites.rhizome.org/anthology/lialina.html>, consulté le 30/01/23).

458 N. Katherine Hayles, *op. cit.*, p. 290.

### ***Un contrepoint audiovisuel***

La bande sonore est elle aussi générative, puisqu'elle se fabrique avec la voix, le son ambiant des séquences (qui n'est pas systématiquement joué), et les musiques choisies par le programme. À ce titre, on pourrait la comparer à la pratique radiophonique du Hörspiel (« Jeu pour l'oreille »), sorte de pièce sonore hybride à plusieurs couches<sup>459</sup>, qui oscille entre narrativité et musique. Celui-ci serait ici revisité avec des strates non plus fixées, mais interchangeableables, et avec une composante vidéo. Dans la plupart des versions de *Soft Cinema*, la voix joue un rôle fondamental, puisqu'elle lie, comme la voix off au cinéma, le déroulement inattendu des séquences. Le décrochage acoustique n'est plus interactif comme dans *The Erl King*, mais déclenché par le programme, selon les règles de l'artiste<sup>460</sup>. Sans le formuler explicitement, Manovich, en faisant cohabiter — tout comme dans le cinéma en split-screen — des séquences audio-vidéo spatialisées, doit arbitrer quelles vidéos doivent devenir, à des moments donnés, muettes, pour conserver une cohérence. Lev Manovich, sur cet aspect, se réfère à Eisenstein et l'idée d'un contrepoint audiovisuel, employée dans le contexte cinématographique de la fin des années 1930, au sein d'un manifeste dont Eisenstein était l'un des auteurs :

Seule l'utilisation du son en guise de contrepoint vis-à-vis d'un morceau de montage visuel offre de nouvelles possibilités de développer et de perfectionner le montage. Les premières expériences avec le son doivent être dirigées vers sa non-coïncidence avec les images visuelles. [...] La méthode du contrepoint appliquée à la construction du film sonore, non seulement n'altérera pas le caractère international du cinéma, mais rehaussera encore sa signification et son pouvoir de culture à un degré inconnu jusqu'à présent<sup>461</sup>.

---

459 « [...] une pièce radiophonique composée d'éléments sonores hétérogènes. Musique, bruits, voix et silence s'y mélangent pour donner corps à une œuvre destinée aux ondes et dont le caractère expérimental l'affranchit de toute forme de hiérarchisation » (Philippe Baudouin, « Du Hörspiel », revue Syntone, 2011, <http://syntone.fr/du-horspiel/>, consulté le 30/01/23).

460 Manovich se présente aussi dans les crédits comme designer sonore.

461 Sergueï Eisenstein, Vsevolod Poudovkine, Grigori Alexandrov, « Contrepoint orchestral », 1928, in S. M. Eisenstein, *Le film : sa forme, son sens*, Paris, Christian Bourgois, 1976, p. 19-21.

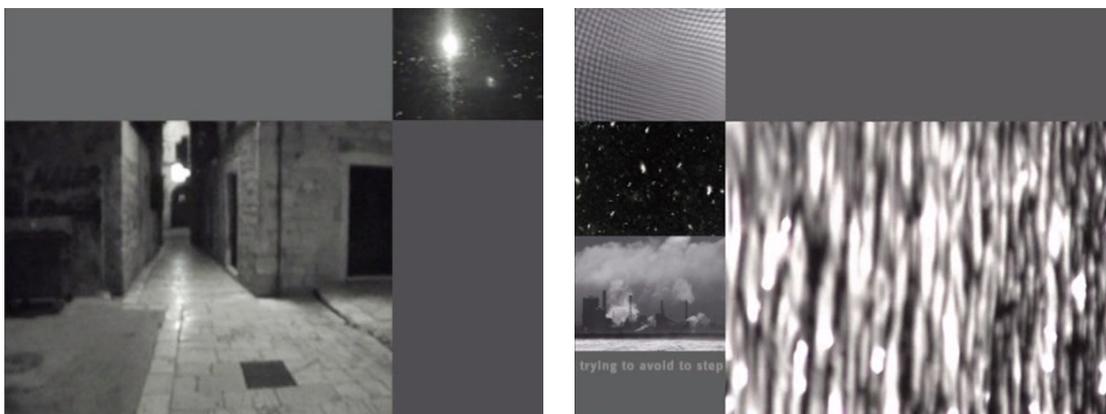
**Absences. Une autre version de *Soft Cinema* qui utilise des sons synthétiques**

Fig. 112 : Andreas Kratky, *Absences* (2004)  
© DVD *Soft Cinema*, MIT Press (2005)

Il existe plusieurs « éditions » de *Soft Cinema*. Chacune d'entre elles utilise une base de données différente, et un système narratif différent. *Texas* (2002) et *Hamburg* (2003) sont basées sur les fictions (*GUI*) de Manovich. *Hamburg* utilise trois différentes voix pour la narration. *Mission to Earth* (2003) a été réalisé à l'occasion d'une sortie DVD. La dernière édition de *Soft Cinema*, *Absences* (2004) a été réalisée par l'artiste et développeur Andreas Kratky<sup>462</sup>. Celui-ci a voulu utiliser le logiciel *Soft Cinema* sans avoir recours aux fictions sonores. Pour lui, la narrativité peut se créer chez l'utilisateur sans histoire préétablie. Il a créé un corpus de séquences tournées dans des environnements naturels, d'autres urbains.

Pour la base de données d'*Absences*, chaque séquence a été analysée en fonction de différents paramètres visuels tels que la luminosité, le contraste, la texture, l'activité, la différence entre le centre et la périphérie et la différence entre la partie supérieure et la partie inférieure. La bande-son d'*Absences* provient du même espace de paramètres : les courbes décrivant l'évolution temporelle de l'activité, les valeurs de luminosité et de contraste des clips individuels ont été utilisées pour synthétiser des sons qui ont ensuite été combinés avec le son du lieu. Le rythme du montage filmique crée également un rythme sonore. La bande sonore globale du film suit la progression des visuels du film, du plus concret au plus abstrait, en passant du son ambiant au son abstrait stratifié [Nous traduisons]<sup>463</sup>.

462 Il est aussi l'auteur de la thèse *Les auras numériques : pour une poétique de la base de données*, op. cit.

463 « For the *Absences* database each film clip was analyzed in regards to several visual qualities such as brightness, contrast, texture, activity, the difference between the center and the periphery, and the

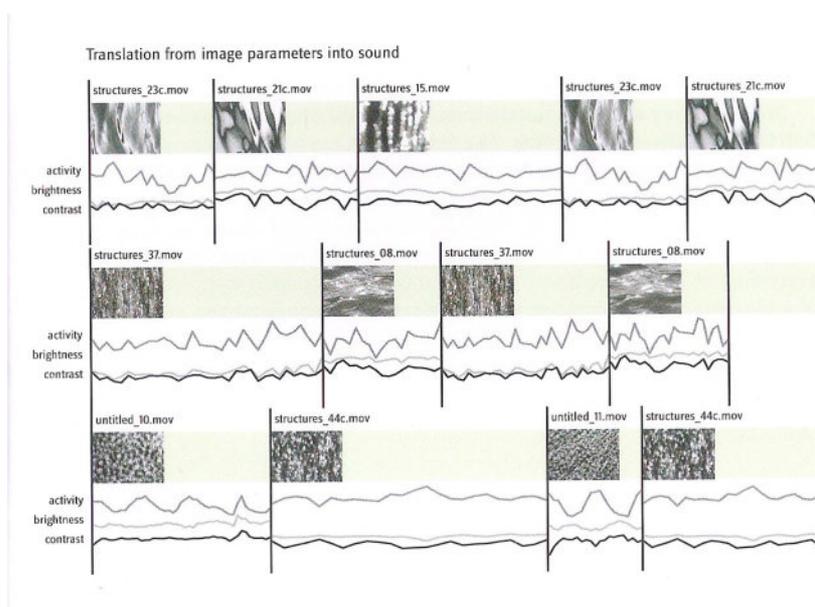


Fig. 113 : Table de correspondances entre séquences vidéo et son  
 © Andreas Kratky, Livret du DVD *Soft Cinema*, MIT Press (2005)

Andreas Kratky fabrique ici un système de correspondance entre mouvement et son, une sonification<sup>464</sup> synthétique des qualités visuelles d'une vidéo. Ces sons synthétiques seront combinés aux sons ambiants et aux autres musiques jouées.

Pour résumer ces relations entre récit et base de données propres à *Soft Cinema*, nous pouvons utiliser les termes de N. Katherine Hayles, qui envisage ces deux architectures comme des « symbiotes naturels<sup>465</sup> », c'est-à-dire pour elle des « organismes d'espèces différentes qui entretiennent une relation mutuellement bénéfique ». Dans ce type de forme hybride, des segments narratifs portent des associations d'images et de sons fabriquées par des mots-clés, et mettent en jeu des sonifications de l'image.

difference between the upper and the lower part. The soundtrack of *Absences* comes from the same parameter space: the curves describing the temporal development of the activity, brightness and contrast values of the individual clips were used to synthesize sounds that were then combined with location sound. The rhythm of the filmic montage also creates a sonic rhythm. The overall film soundtrack follows the progression of film visuals from more concrete to more abstract by moving from location sound to the layered abstract sound. » (Page du site [softcinema.net](http://www.softcinema.net), qui n'est plus en ligne, récupéré par la Wayback Machine de L'Internet Archive : <http://web.archive.org/web/20120229005700/http://www.softcinema.net/editions.htm>, consulté le 30/01/23).

464 Nous avons exploré ce principe dans la première partie, section 1.4.4.

465 N. Katherine Hayles, *Lire et penser en milieux numériques*, op. cit., p. 282.

### 2.3.4 Dé-monter et remonter le cinéma. Les performances audiovisuelles de SLIDERS\_lab

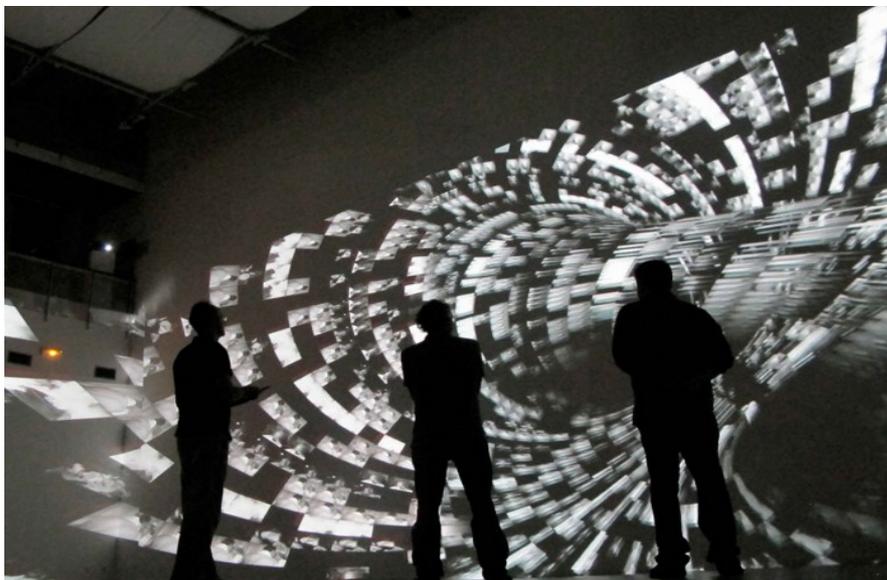


Fig. 114 : SLIDERS\_lab, *Sky Memory Project*, 2011-2014  
© Jean-Marie Dallet / SLIDERS\_lab

Dans les deux pièces étudiées dans les précédentes sections, *The Erl King* et *Soft Cinema*, l'artiste agit comme un préparateur. Il réalise les médias, les indexe dans une base de données, écrit un programme chargé de les agréger selon ses règles. Contrairement à un montage linéaire et fixé, des moments d'ouverture potentielle y sont définis puis c'est un opérateur autre qui va monter une version dans des conditions de temps réel. Dans le cas de *The Erl King*, l'opérateur est un participant du public (il y a interactivité), dans *Soft Cinema* l'opérateur est la machine elle-même (il y a générativité). Dans d'autres cas, c'est l'artiste lui-même qui va manipuler son programme, et donc sa base de données, en direct. Habituellement caractérisées en anglais par l'expression *Live Cinema*<sup>466</sup> ou *Live A/V* — pour *Live Audiovisual Performance* —, les formes performatives de l'image-son sont souvent apparentées à des formes étendues de concerts intégrant de l'image projetée. Leur généalogie est en réalité bien plus mul-

466 L'artiste Mia Makela définit le *Live Cinema* comme une pratique intrinsèquement artistique de « création simultanée du son et de l'image en temps réel, par des artistes sonores et visuels qui collaborent à hauteur égale sur des concepts élaborés ». Ces performances, plus proches de la poésie ou de la musique que du cinéma narratif, ne sont pas présentées dans des clubs, comme pour les DJ et les VJ, mais dans des musées, des théâtres ou des salles dédiées (Mia Makela, « On l'appelle le Live Cinema », in MCD — Magazine des Cultures Digitales, hors-série, *Live A/V. Performances audiovisuelles*, avril 2010, p. 6).

tipte. S’y entrecroisent des hybridations de musique visuelle, de performance multimédia, de cinéma élargi et d’art vidéo, couplées aux pratiques de *Djing* et de *Vjing*<sup>467</sup> qui ont émergé avec la musique électronique. Ces pratiques *live*, dont les formes intègrent une part numérique, sous-entendent la présence de systèmes programmés, et/ou programmables, qui permettent à l’artiste d’intervenir en temps réel sur des paramètres de l’œuvre : son déroulement, voire même l’espace dans lequel elle se développe.

Dans ce contexte, on peut observer dans les performances de SLIDERS\_lab<sup>468</sup> une forme de synthèse des possibilités numériques appliquées à une forme ouverte de cinéma. Dans un contexte de performance, quand le corpus préparé devient important, il nécessite de prendre en compte la recherche dans la base de données, mais aussi de mettre en scène la base de données elle-même. Comme le formule Jean Marie Dallet, « Cela doit se faire par la construction d’architectures d’images, de sons, de textes, capables de proposer des modèles de circulation dans l’espace des données<sup>469</sup> ».

### Dé-monter le film



Fig. 115 : SLIDERS\_lab, performance Résonances  
IRCAM, Paris, France, 2006  
© SLIDERS\_lab



Fig. 116 : SLIDERS\_lab, performance  
IMAL Bruxelles, Belgique, 2008  
© SLIDERS\_lab

Dans *Sliders\_1* (2005-2009)<sup>470</sup>, performance de cinéma interactif, un film est fabriqué en direct et en public par trois performeurs. Les images et les sons proviennent du film

467 Ou *Video Jockey*, l'équivalent visuel des DJ. Souvent en retrait, les VJs mixent des images sur un écran situé derrière le DJ, qui lui se trouve généralement au premier plan.

468 Les deux membres de SLIDERS\_lab sont Jean-Marie Dallet et Frédéric Curien.

469 Jean-Marie Dallet, « Figures de l'interactivité », In *Anomalie\_digital art n° 3*, Clermont-Ferrand, Un, Deux... Quatre Éditions, mars 2003, p. 35.

470 Plusieurs captations vidéo sont disponibles en ligne sur le Vimeo de SLIDERS\_lab, par exemple le set 5, <https://vimeo.com/144611367>, ou le set 8, <https://vimeo.com/144611371>, consultés le 30/01/23.

*Psychose* (1960) d'Alfred Hitchcock, et de son remake en 1998, par Gus Van Sant. Trois espaces indépendants communiquent entre eux pendant la performance. Deux espaces sont liés aux séquences vidéo : l'un gère la base de données, l'autre permet la manipulation des séquences. Il est possible, à cet endroit, pour le performer, d'étirer ou de fractionner les séquences. Le troisième et dernier espace est dédié à la sélection et à la jouabilité des sons.

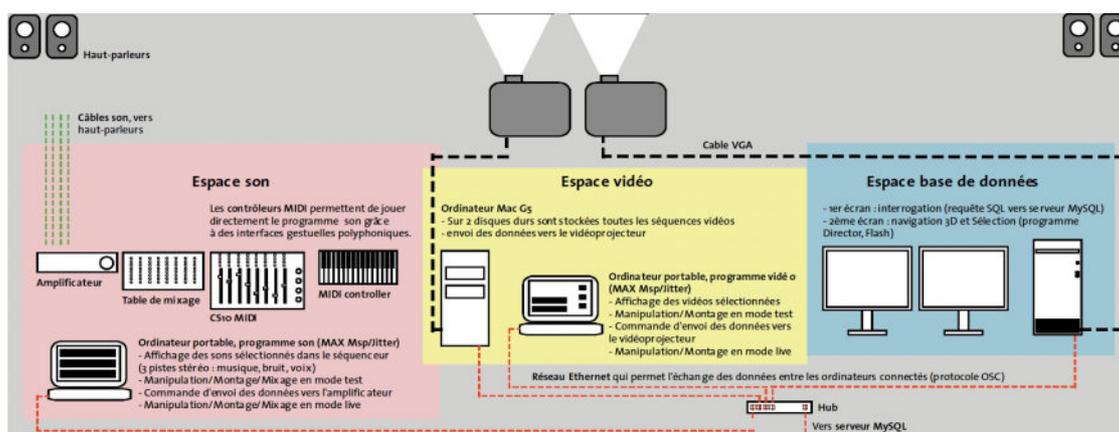


Fig. 117 : SLIDERS\_lab, documentation du dispositif  
© SLIDERS\_lab

Les sons ont été catégorisés de cette manière :

La philosophie générale qui a prévalu au développement du programme « Sons » a permis la distinction dans l'univers sonore de trois plans, ceux de la musique, du bruit et des voix. Trois listes ont donc été établies, elles-mêmes subdivisées en sous-listes, par exemple celle des « Coups de feu » ou des « Portières qui claquent » qui appartiennent à la liste principale des « Bruits ». Ces listes sont interactives, bien sûr, et permettent ainsi d'atteindre rapidement les éléments à jouer durant la performance<sup>471</sup>.

Chaque séquence vidéo est indexée selon 22 critères, parmi eux : le mouvement de caméra<sup>472</sup>, le sens du zoom s'il existe, la valeur de plan, le point de vue caméra. Certaines natures binaires de plan (une porte qui s'ouvre ou se ferme, une lumière qui s'allume

471 Jean-Marie Dallet, « Un cinéma de la base de données », in Caroline Renard (dir.), *Images numériques ?*, op. cit., p. 231.

472 Sur un principe similaire à celui employé par Lev Manovich et d'Andreas Kratky dans *Soft Cinema*, voir à la section précédente 2.3.3.

ou s'éteint) sont décrites comme des interrupteurs (0 ou 1). Environ 600 séquences ont été indexées grâce à ce principe, et 1500 sons ont été décrits. Grâce à toutes ces métadonnées, il est possible d'appeler une séquence grâce à des requêtes, ensembles de mots-clés tels que « gros plan, Marion, rouge ». Ces séquences sont ensuite déposées, à la manière d'une playlist ou d'une file d'attente informatique, dans ce que les artistes appellent un « réservoir<sup>473</sup> ».

Le performeur peut intervenir en temps réel sur la microstructure d'un plan visuel ou d'un son, mais aussi sur la macrostructure globale du film. Si les appairages originaux entre image et son restent conservés dans la base de données, il est aussi possible de travailler de manière indépendante sur les deux composantes.

L'objectif du cahier des charges établi pour le son était de fournir un instrument jouable, modulaire et ouvert, permettant aux performeurs de contrôler en direct l'ensemble des paramètres sonores habituellement accessibles dans le processus de postproduction audio au cinéma. Montage, mixage et traitement spatial du son à l'image devaient être rendus manipulables en direct. Dans la tradition de la sonorisation du film muet, les composantes de la bande son, musique, bruits et voix, peuvent ainsi être librement séparées du continuum filmique pour être ensuite réagencées et séquencées à volonté dans l'idée du SLIDERS\_lab, soit pour accompagner les séquences de manière « réaliste », soit pour les dissocier des images projetées et créer un sens externe autonome<sup>474</sup>.

Il y a un transfert de ce qui se fait habituellement en production par des techniciens, à l'artiste performeur sur scène. Comme le formule la professeure et théoricienne Martine Beugnet au sujet des œuvres numériques de réemploi, les actions consistant à extraire et à remonter sont fondamentales. Le montage non linéaire offre, selon elle, ouvre de nouvelles perspectives de « malléabilité de la bande-son<sup>475</sup> » et de synchronisation audiovisuelle. C'est bien le cas ici, chez les SLIDERS\_Lab.

---

473 *Ibid.*

474 Frédéric Curien et Jean-Marie Dallet, « Donner forme aux mémoires », in Jean-Marie Dallet (dir.), *Cinéma, interactivité et société, op. cit.*, 2013, p. 400.

475 Martine Beugnet, *Le cinéma et ses doubles. L'image de film à l'ère du foundfootage numérique et des écrans de poche*, Lormont, Le bord de l'eau, 2021, p. 25.

### ***Le son des images démultipliées***



Fig. 118 : SLIDERS\_lab, *TMWKTM* (2009-2015)  
© sliderslab.com

Une autre opération de démontage du cinéma des SLIDERS\_lab met en jeu les relations de l'image et du son, cette fois-ci de manière générative. Dans *TMWKTM* (2009-2015)<sup>476</sup>, un autre film d'Alfred Hitchcock — *The Man Who Knew Too Much* (1936) — a été décomposé en de multiples séquences. Séparées les unes des autres et spatialisées en profondeur, elles vont venir dans des mouvements continus s'approcher ou s'éloigner de l'écran, jusqu'à disparaître lorsque leur surface devient trop importante ou trop réduite. Il s'agit, comme le *Soft Cinema* de Lev Manovich, d'un film automatique. La particularité de *TMWKTM*, au-delà de sa forme singulière, est le traitement de la partie sonore.

Une caméra virtuelle va alors automatiquement entraîner le spectateur à regarder le film, de loin comme une carte, de plus près comme un territoire, sous de multiples angles de vue, avec inlassablement un jeu de prises de son en temps réel effectué à l'aide de micros. Les fragments sont en fait des sphères de couleur rouge, verte ou bleue qui touchent les séquences d'images et en prélèvent, l'instant de ce contact, la rumeur, c'est-à-dire les voix, les bruits et les musiques comme autant de pistes séparées. Chaque fragment du film démonté constitue la part d'une archive qui s'anime au contact des têtes de lecture sonores<sup>477</sup>.

476 Une captation vidéo est disponible en ligne sur le site de SLIDERS\_lab, [http://www.sliderslab.com/pages\\_fr/TRAVAUX/TMWKTM.html](http://www.sliderslab.com/pages_fr/TRAVAUX/TMWKTM.html), consulté le 30/01/23.

477 Christophe Domino, Jean-Marie Dallet, « Une archive active pour l'Institut pour l'abolition de la guerre. Principes pour une collaboration », *op. cit.*

Plutôt que de laisser les images et les sons de chaque séquence se jouer (les vidéos pourraient toutes se jouer et le volume sonore de chaque vidéo aurait pu se régler automatiquement en fonction de son rapprochement par exemple), des sphères colorées se promènent à l'écran et jouent le rôle d'une tête de lecture sonore. Au lieu de cela, toutes les scènes qui flottent dans l'espace virtuel sont en pause. Lorsque l'une des boules entre en collision avec une scène, alors la vidéo (l'image et le son) se joue. En complément, la vidéo prend la couleur de la boule, de manière à ce que le visiteur perçoive quelles sont les séquences en lecture. Dans cet espace, la multiplicité des images contraint l'œil à travailler plus que l'oreille, aussi les segments sonores déclenchés prennent une importance particulière. Ils créent des moments de focalisation.

### ***Le film comme base de données***



Fig. 119 : SLIDERS\_lab, *Sky Memory Project*, 2011-2014  
© Jean-Marie Dallet / SLIDERS\_lab

Dans une autre de leurs performances interactives, *Sky memory Project* (2011-2014)<sup>478</sup>, les artistes proposent au public d'assister à une navigation performée dans des architectures plastiques en deux ou trois dimensions (des tores, des hélices...). Les surfaces de ces dernières accueillent des images, tandis que des sons évoluent, formant un ensemble polyphonique et dynamique de bruits, de musique et de voix. Les séquences utilisées relèvent du réemploi, elles sont tirées de différents films d'aventure. Un ou plusieurs opérateurs, devant l'écran, et devant le public, manipulent une tablette de contrôle.

<sup>478</sup> Une captation vidéo est disponible en ligne sur le site de SLIDERS\_lab, [http://www.sliderslab.com/pages\\_fr/TRAVAUX/Sky\\_Memory\\_Project.html](http://www.sliderslab.com/pages_fr/TRAVAUX/Sky_Memory_Project.html), consulté le 30/01/23.

La base de données est en quelque sorte rendue tangible, mise en scène et mise en flux, dans ses composantes sonores et visuelles.

À la piste son et à la piste image du cinéma, il faut désormais en adjoindre une autre, celle du programme. Celle-ci crée entre les images disséminées dans la mémoire de l'ordinateur des points de tension et des contacts parfois insensibles, des passages initiés par des précurseurs sombres qui tracent à l'avance une voie, des lignes, des trajectoires et des parcours, qui resteront incertains et transitoires, qui s'actualiseront puis s'effaceront. Un espace sourd, antérieur à toute trace spatiale<sup>479</sup>.

Elle constitue, avec le programme qui la matérialise, le cœur de l'œuvre.

### 2.3.5 Le son des *Vlogs* et l'analyse automatique. *Sound Stories*, de Christian Marclay



Fig. 120 : Christian Marclay, *The Organ* (2019)  
© ARTnews / Benoît Florençon

#### ***Un travail de réemploi***

Une grande partie du travail de l'artiste Christian Marclay, de sa pratique de « Turntablism<sup>480</sup> » expérimental jusqu'à son travail de montage vidéo, se fonde sur le réemploi. Son travail vidéo est généralement diffusé sous forme d'installations muséales : des montages s'appuyant sur des fragments de films hollywoodiens, comme

<sup>479</sup> Jean-Marie Dallet, « Un cinéma de la base de données », *op. cit.* p. 229.

<sup>480</sup> Souvent des sets d'improvisation sur des platines, les vinyles pouvant être rayés ou préalablement détériorés.

dans *Telephones* (1995), *Video Quartet* (2002), ou bien le célèbre *The Clock*<sup>481</sup> (2010). Si les rouages numériques sont a priori étrangers à sa pratique, il y a déjà dans un travail monumental comme celui effectué pour *The Clock* une confrontation à des méthodes de recherche qui s'appuient sur la culture numérique.

Pour *The Clock*, j'ai employé des gens qui regardaient des films toute la journée. Grâce aux technologies d'aujourd'hui, nous irions sans doute beaucoup plus vite. Cela dit, nous avons déjà beaucoup utilisé Internet pour chercher des références. Nous consultions des forums sur lesquels des obsédés des montres savent exactement à quel moment et dans quel film de James Bond se trouve telle ou telle indication horaire<sup>482</sup>.

Lorsque l'entreprise Snap Inc. — à l'origine du réseau social Snapchat — lui propose en 2018 de réaliser pour le festival d'art Lions à Cannes une œuvre similaire à *The Clock*, (qui utiliserait les technologies et le corpus de Snapchat), il refuse au départ la collaboration. Finalement, sa curiosité pour des technologies qui lui sont inconnues et la perspective de liberté artistique qui lui est offerte prend le pas. Le projet évolue au fil d'une collaboration de près d'une année et se conclut par la création de cinq installations numériques<sup>483</sup>, chacune investissant un espace dédié. L'ensemble constitue l'exposition *Sound Stories* (2018).

### **Le son des « Snaps »**

Il est proposé à Marclay de travailler avec le corpus de Snapchat, composé de *Snaps*. Les *Snaps* sont des séquences vidéo de 10 secondes postées par des utilisateurs, conçues pour être éphémères et seulement consultables par son réseau. Une fois consultées par leur destinataire elles sont supprimées, sauf si elles sont partagées ou rendues publiques, auquel cas elles peuvent rester en ligne plus longtemps et vivre une seconde vie. Ce sont, à l'époque de leur création en 2011, les premières *Stories*, format

481 Pour *The Clock*, Christian Marclay et son équipe ont créé un montage de 24 h de séquences préexistantes, empruntés à des films hollywoodiens. Chacune des scènes fait apparaître une montre, une horloge, ou tout autre objet qui indique l'heure, celle-ci devant correspondre au temps du film écoulé. Le film est projeté de manière à ce qu'un visiteur qui rentre dans la salle à 14 h verra les scènes qui correspondent à cette même heure.

482 « Christian Marclay : "La musique m'a sorti de mon atelier" », entretien avec Emmanuel Grandjean, *The Art Newspaper*, janvier 2021, <https://www.artnewspaper.fr/interview/christian-marclay-la-musique-m-a-sorti-de-mon-atelier>, consulté le 30/01/23.

483 Dans la villa du centre d'art La Malmaison à Cannes pour la première, puis l'exposition tournera dans différents lieux.

repris depuis par la plupart des autres réseaux sociaux. Le cadre de cette collaboration avec Span Inc. permet à Marclay d'avoir un accès à près de deux cents millions de Snaps, ce qui correspond à la partie « Our Story » de Snapchat, qui regroupe des vidéos publiques autour de points communs (notamment la localisation). Au-delà de la base de données, Marclay s'inscrit aussi ici dans des pratiques d'artistes qui utilisent des CGU (Contenus Générés par les Utilisateurs), contenus au statut encore indéterminé<sup>484</sup>.

La communication de Snapchat est essentiellement axée sur l'image<sup>485</sup>. Prenant le contrepied du discours de l'entreprise, Marclay axe toute sa réflexion sur le sonore.

Je voulais remettre le son au premier plan, pour ne plus seulement regarder l'image, mais pour considérer le téléphone mobile comme un microphone plutôt qu'une caméra [Nous traduisons]<sup>486</sup>.

Pour lui, les utilisateurs des réseaux sociaux sont en effet souvent attentifs au cadrage, à l'arrière-plan, mais ne pensent pas aux sons qu'ils postent, qui sont pourtant tout aussi porteurs de sens, et de surcroît analysables par des programmes.

Le titre de l'exposition, *Sound Stories*, témoigne du point de vue adopté par Marclay. Les cinq installations audio-vidéo qui constituent l'exposition sont *All Together*, *Tinsel Loop*, *The Organ*, *Talk to Me/Sing to Me* et *Sound Tracks*. Toutes ces installations sont exposées dans la pénombre. Parmi elles, deux installations sont interactives (*The Organ* et *Talk to Me/Sing to Me*), ce qui est inédit chez Marclay. Plus encore, la plupart de ces installations ont recours à une technologie d'analyse et d'indexation fréquentielle du son.

484 Le groupe de recherche « Après les réseaux sociaux », initié par Gala Hernández López et Allan Deneuve travaille sur ce type de travaux, <http://after-social-networks.com/>, consulté le 30/01/23.

485 Notamment sur la caméra qui s'active dès le lancement de l'application et qui permet de publier du quotidien en instantané.

486 « I wanted to bring the sound to the forefront, not just to look at the image, but to think of the phone as a microphone rather than a camera » (« Conversation. Andrew Lin, Christian Marclay, and Tim Marlow », Cannes, juin 2018, in Rita Gonzales [dir.], *Christian Marclay. Sound Stories*, catalogue d'exposition, Santa Monica, LACMA/Snap Inc., 2019, p. 83.

### ***Une indexation automatique par la sonorité***

Lorsque l'équipe d'ingénieurs a permis, dans les premiers temps de la collaboration, à Marcaly de se familiariser avec le moteur de recherche, l'artiste a très vite proposé aux ingénieurs de mettre en place une fonctionnalité permettant de rechercher des vidéos par les fréquences dominantes qui y sont détectées. Les personnes qui ont l'oreille absolue le savent : elles peuvent reconnaître, si elles entendent un train qui passe, la note produite ou tout au moins savoir entre quelles notes se situe la fréquence entendue. Le « La » du diapason est par exemple à la fréquence de 440 Hz. Plus le son est complexe, plus il contient de fréquences différentes (les partiels), mais si l'on considère qu'en parlant, nous émettons une sorte de polyphonie de fréquences, il est possible techniquement d'analyser, par ce que l'on nomme les transformées de Fourier, la répartition des fréquences produites, et donc de déterminer celles qui sont les principales émises. Il devient possible de classer les séquences vidéo grâce à l'analyse de leur son, par note.

### ***All Together, composition polyphonique pour smartphone***



Fig. 121 : Christian Marcaly, *All Together* (installation de *Sound Stories*)  
Vue de l'installation au LACMA, Los Angeles, 2019  
© Christian Marcaly / LACMA

La première installation du parcours, *All Together*, est la pièce qui se rapproche le plus des pièces vidéo de *Found Footage* de Marcaly, telle une introduction-amorce, un accordage sur ce qui va suivre. Il s'agit d'un travail de composition, montage multiécran spatialisé sur 10 iPhone positionnés en ligne. La base de données est un corpus de

400 séquences (des *Snaps*) sélectionnées par Marclay. L'image et le son de chaque séquence restent couplés dans la composition. Le son est diffusé par les haut-parleurs de l'iPhone, l'image par son écran. La volonté de Marclay est de rester fidèle à l'appareil qui a permis de capturer les séquences originales, aussi les sons graves sont peu audibles et l'image à son échelle naturelle. Ni amplification, ni projection, un mini-opéra polyphonique de 4 minutes 30, joué en boucle, se regarde et s'écoute. Si le montage semble manuel, puisqu'il reste le même et est joué en boucle, les microvidéos (une dizaine de secondes chacune) ont été réunies dans le montage car elles donnent à entendre des bandes-son similaires. On entend dans l'installation des cliquetis, des aboiements, de la pluie, des percussions. Des vidéos qui visuellement n'ont rien à voir sont parfois jouées ensemble car la bande-son de l'une — le bruit de l'huile sur le feu — s'apparente à celui d'une autre — celui d'une machine à coudre. Parfois le visuel aussi est similaire, toutes les vidéos vont par exemple montrer une personne (pour chacune différente), qui marche. L'ensemble est très poétique, et donne la sensation de « glisser » de vidéo en vidéo par le fil sonore (grincements, boîtes à musique, flûtes, sirènes, sifflements...) tissé par l'artiste.

### ***Tinsel Loop, une boîte à musique audio-visuelle***

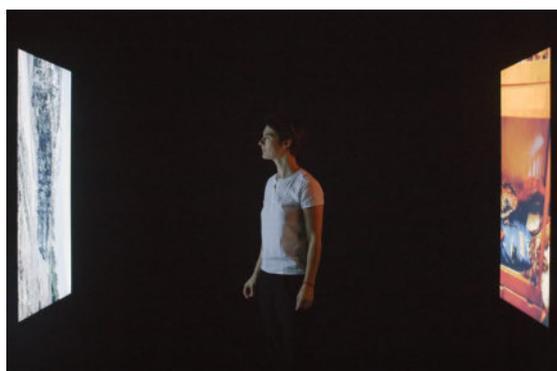


Fig. 122 : Christian Marclay, *Tinsel Loop* (installation de *Sound Stories*)  
 Vue de l'installation au LACMA, Los Angeles, 2019  
 © Christian Marclay / LACMA

En 2005 Marclay avait composé pour une boîte à musique une mélodie atonale, *Tinsel*. Le mécanisme était placé dans une boîte sur laquelle il était écrit « Silent » ; si on ouvrait la boîte, on pouvait lire son anagramme « Listen ». Dans *Tinsel Loop*, la même

mélodie jouée en boucle est réinterprétée, non plus par une boîte à musique, mais par la succession des micro bandes sonores des séquences qui viennent, sur deux écrans placés l'un en face de l'autre, s'afficher en fonction de la note jouée, toujours grâce au système d'analyse fréquentielle mis au point par Marclay et l'équipe de Span Inc. Parce que celui-ci est utilisé cette fois pour détecter des sons plus musicaux que pour l'installation précédente — *All Together* —, nous pouvons remarquer qu'il s'avère très similaire à celui utilisé par l'artiste Cory Arcangel, dans son travail de *supercut*<sup>487</sup> sur des vidéos de YouTube<sup>488</sup>. En somme, à la place d'une des notes de la mélodie est jouée une vidéo dont la sonorité générale s'approche de la note.

### ***The Organ. L'orgue sonore à images***



Fig. 123 : Christian Marclay, *The Organ* (installation de *Sound Stories*)

Vue de l'installation au LACMA, Los Angeles, 2019

© Christian Marclay / LACMA

487 Montage d'images trouvées (*Found Footage*) autour d'un thème particulier.

488 Dans l'œuvre *A Couple Thousand Short Films About Glenn Gould* (2006), Cory Arcangel reconstitue, à l'aide de notes prélevées sur des vidéos publiées sur YouTube, la *Variation n° 1* de Jean-Sébastien Bach, première des *Variations Goldberg* (circa. 1740). La particularité de cette variation est qu'elle ne comporte que des notes jouées successivement, ce qui se prête bien au jeu posé par Arcangel. Le Titre de la pièce d'Arcangel fait référence à un film de François Girard, *Thirty Two Short Films About Glenn Gould* (1993) qui est composé de 32 séquences (dont le générique) autour de Glenn Gould. On connaît par ailleurs les méthodes de montage du pianiste, qui, pointilleux, montait en studio ses interprétations parfois presque note à note, à partir de multiples prises. La méthode de montage d'Arcangel est la suivante. Il télécharge au préalable un grand nombre de vidéos YouTube (musiciens, chats qui jouent de la musique...), puis isole à l'oreille des séquences courtes ne comprenant qu'une note. Ces séquences sont entrées dans une base de données, puis traitées par un logiciel. Les logiciels de montage de l'époque ne lui permettant pas d'atteindre la vitesse souhaitée, il a développé le sien, qu'il a nommé *Gould Pro*. Le film — un diptyque — utilise au final 1106 séquences, chacune étant une note isolée couplée à son visuel vidéo original. À partir de cette discrétisation, Arcangel reconstitue la mélodie note par note (vidéo par vidéo) pour recréer un flux artificiel, sorte de montage en *supercut*.

Dans *The Organ* (2018)<sup>489</sup> le dispositif technologique de recherche par fréquence est mis à disposition de l'utilisateur par l'intermédiaire d'un clavier musical comportant les touches blanches et noires habituelles. En jouant une composition, on ne produit pas directement du son, mais pour chaque note s'affiche une colonne de vidéos (des *Snaps*) qui ont été sélectionnés par similarité de fréquence avec la note potentielle que l'on souhaite jouer. Si le visiteur appuie sur la touche de la note « La » au milieu du clavier, le moteur de recherche va chercher les vidéos dont le son général se rapproche le plus possible d'un « La » et va jouer une colonne de vidéos simultanées au même niveau que la touche, devant le visiteur. Le résultat sonore, pour une seule touche « La », est donc une superposition sonore de « La » approximatifs — assez proche de ce que l'on appelle un « cluster » en musique contemporaine<sup>490</sup> — puisque les sons des séquences ne sont pas ce que l'on appelle des sons purs<sup>491</sup>. Il est a priori possible de jouer n'importe quelle composition, et donc possible de jouer des accords, c'est-à-dire d'enfoncer plusieurs touches, ce qui génère alors plusieurs colonnes vidéo, et donc autant de clusters. Marclay retrouve ici, à partir de fragments numériques indexés à partir de leurs caractéristiques spectrales, un son « sale » dont il est familier.

### ***L'écho des grottes numériques***

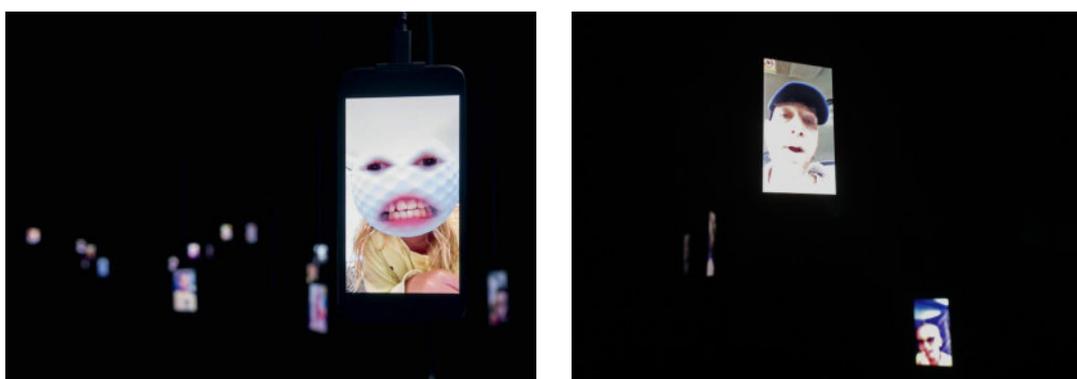


Fig. 124 : Christian Marclay, *Talk to Me / Sing to Me* (installation de *Sound Stories*)  
Vue de l'installation au LACMA, Los Angeles, 2019

489 Une captation vidéo est disponible en ligne sur le Facebook du LACMA, <https://www.facebook.com/watch/?v=551036715723791>, consulté le 30/01/23.

490 Des « grappes » de sons très serrés, en accords compacts, formant généralement un total chromatique. On emploie fréquemment le terme de cluster à propos de groupes de sons produits en frappant le clavier du piano avec le plat de la main, le poing ou l'avant-bras [...]. Un des premiers compositeurs à avoir exploité cette technique de jeu est Henry Cowell » (Jean-Yves Bosseur, *Vocabulaire de la musique contemporaine*, Paris, Minerve, 2013, p. 29).

491 Ou sons sinusoïdaux : qui ne sont composés que d'une seule fréquence.

© Christian Marclay / LACMA

Dans *Talk to Me/Sing to Me*, des smartphones sont suspendus devant les visiteurs. Sans intervention particulière, ces derniers affichent les phrases « Talk to Me » ou « Sing to Me ». Si l'on chante, parle ou crie, les smartphones vont, à la manière d'un écho, jouer des séquences vidéo dont les voix vont reconstituer par imitation ce qui a été dit. Le système informatique utilise d'une part la reconnaissance fréquentielle pour aller chercher des vidéos, mais également un système de transformation sonore pour forcer les bandes sonores à s'aligner sur la génération de voix. Nous étudierons, dans la troisième partie, des systèmes d'analyse de la voix assez proches de cet exemple.

### ***Le ralenti de la déconstruction***

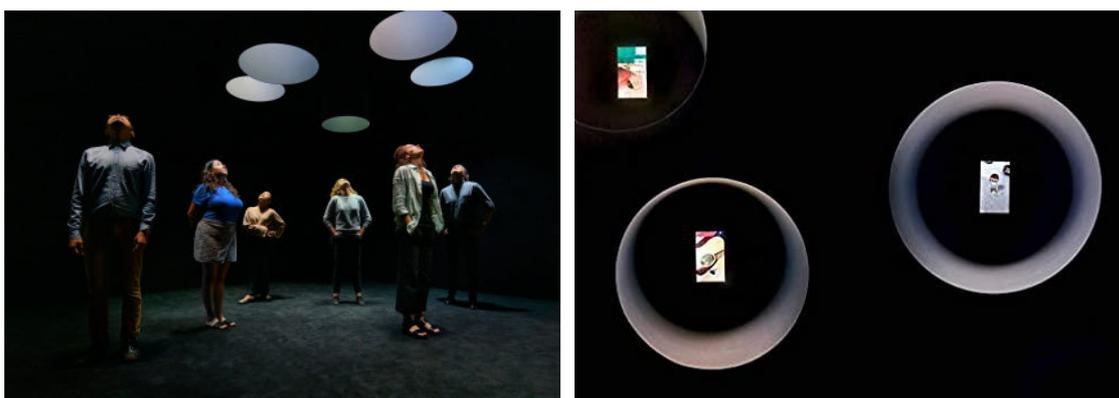


Fig. 125 : Christian Marclay, *Sound Tracks* (installation de *Sound Stories*)

Vue de l'installation au LACMA, Los Angeles, 2019

© Christian Marclay / LACMA

La pièce conclusive, quant à elle, s'intitule *Sound Tracks*. Plus dérangement, elle n'offre pas d'emblée de visuels. Des sons indistincts, bourdonnements inquiétants, sont diffusés dans l'espace. Le visiteur peut, s'il lève la tête, se rendre compte que des tablettes ont été placées au plafond. Chacune d'entre elles diffuse une scène issue de Snaps ralentie trois ou quatre fois, le son également. Marclay utilise un effet disponible sur Snapchat, mais très peu utilisé : l'effet *Turtle* (Tortue). On peut alors comprendre, et associer la scène jouée au son. Marclay joue sur le fait connu que l'être humain

cherche en général toujours à identifier la source d'un son et sépare spatialement le son de l'image.

Il n'existe pas de réseau social exclusivement sonore, mais ici Marclay replace le son au premier plan en explorant l'aspect musical des bandes-son de *Vlogs*. En montrant que le son numérisé est tout aussi analysable que l'image, il pose également la question de la reconnaissance vocale et des écoutes, que nous traiterons dans la troisième partie.

## **2.4 Les images et les sons des paysages technologiques**

Nous avons, dans le chapitre précédent, évoqué des formes artistiques qui s'appuient sur des bases de données nourries de séquences vidéos et de sons. Dans certains cas, comme celui du *Soft Cinema* de Lev Manovich, les films tournés par l'artiste sont indexés (parmi d'autres critères) selon leur localisation. On parle alors de médias localisés, ou situés<sup>492</sup>. Le programme peut alors, lorsque cela est pertinent vis-à-vis des intentions de l'artiste, charger des vidéos dont le lieu de tournage est proche géographiquement. Avec l'avènement de nos appareils numériques, une grande partie des images et des sons capturés dans nos environnements sont d'ailleurs désormais automatiquement localisés. De la webcam située aux hypermédias cartographiques, ce chapitre explore la relation entre les images et les sons au sein de « paysages technologiques<sup>493</sup> » pionniers ou contemporains.

### **2.4.1 Images et sons du monde en direct : du voyage imaginaire à la surveillance**

Le développement des caméras fixes, situées et connectées a progressivement donné aux internautes un accès direct aux images du monde. Plus encadrées actuellement qu'aux débuts du net — elles sont majoritairement produites par les offices de tou-

---

492 Bernard Guelton, « Introduction », in Bernard Guelton (dir.), *Médias situés et mobilités partagées*, Paris, Hermann, 2019, p. 16.

493 Nous utilisons le terme « paysage technologique » en référence à une recherche dirigée par Andrea Urlberger au milieu des années 2000, autour du GPS : « Paysage technologique — théories et pratiques autour du Global Positioning System », [http://www.ciren.org/ciren/laboratoires/Paysage\\_Technologique/index.html](http://www.ciren.org/ciren/laboratoires/Paysage_Technologique/index.html), consulté le 30/01/23.

risme —, ces webcams montrent principalement des monuments ou des bords de plage, et procèdent d'un caractère fonctionnel : rendre compte de la météo, prévenir les files d'attente... Pour ne pas être assimilées à de la surveillance, les images doivent être conformes à un certain nombre de règles, la principale étant que les personnes ne doivent pas être reconnaissables. Le plan large est donc de mise, ce qui confère à l'image une nature particulière, distante.

La webcam apparaît au début des années 1990, avec la naissance du web, et ne sert à l'origine qu'à surveiller le remplissage d'une cafetière. De manière semblable à ce qui s'est produit avec les caméras portables et l'art vidéo, cet appareil simple d'utilisation va très vite être dirigé par des personnes vers eux-mêmes, vers leurs espaces de vie quotidiens, mais de manière amateur, par des anonymes sans revendication artistique. En avril 1996, l'étudiante américaine Jennifer Ringley installe une web caméra et commence à se filmer quotidiennement. Ni les scènes intimes ni les moments d'absence (nombreux et longs car la jeune fille était étudiante) ne sont coupés, ce qui crée une addiction publique immédiate. Le journal filmé se démocratise, c'est le début des réseaux sociaux.

À l'origine, les images envoyées par les webcams sont fixes, envoyées une à une avec un rafraîchissement de plusieurs secondes<sup>494</sup> : une sorte de cinéma, muet, dont le rythme des images est plus lent que celui des premiers temps du cinéma. La nouveauté est que ce cinéma « pauvre », est en direct et accessible à tous sur un navigateur. L'augmentation de la bande passante fera progresser la cadence de rafraîchissement, pour que se fabrique en direct, des séquences d'images de moins en moins saccadées, proches de l'objet film. Le sonore n'est toutefois encore pas un sujet direct, puisqu'il nécessiterait trop de bande passante pour être transmis en direct. Et contrairement à l'image, il n'est pas décomposable en unités discrètes perceptibles.

---

494 Anne Cauquelin, *L'exposition de soi. Du journal intime aux Webcams*, Paris, Eshel, 2003, p. 63.



Fig. 126 : Natalie Bookchin, *Location Insecure* (2005-2007)  
© Vimeo Natalie Bookchin

D'autres webcams, fixes, sont dirigées vers l'extérieur. Elles filment en permanence un environnement extérieur. À l'émergence du Web, de nombreux artistes ont travaillé avec ces nouvelles images, capturées et diffusées en direct. Dans les années 2000, l'artiste Natalie Bookchin<sup>495</sup> a exploité les failles des moteurs de recherche de l'époque dans le but de faire émerger des URL<sup>496</sup> de webcam supposées privées. Elle a ainsi récupéré des images de vidéosurveillance qui n'auraient pas dû être rendues accessibles au public. Le titre de l'une de ces vidéos, *Location Insecure* (2005-2007)<sup>497</sup>, fait par exemple référence à un terme qu'elle intégrait dans sa recherche<sup>498</sup> pour trouver ces images. Elle n'utilise pas les images en direct, mais les récupère et les monte.

Dans *Round the World* (2007)<sup>499</sup>, elle utilise l'un des premiers enregistrements sonores d'Edison sur son phonographe. Celui-ci, intitulé *Around the World on the Phonograph* — qui date d'octobre 1888 —, nous donne à entendre l'inventeur décrire un voyage imaginaire en bateau à vapeur. Dans le film de Bookchin, cette voix devient la voix off d'images de webcam de surveillance provenant de différents endroits du globe.

495 Nous parlerons de ses travaux plus récents, qui utilisent des *Vlogs*, de manière détaillée dans la troisième partie, section 3.4.1.

496 Uniform Resource Locator. Adresse d'une page web à indiquer au navigateur pour qu'il s'y connecte. Ces adresses, quand elles sont rendues publiques, permettent aux moteurs de recherche d'indexer les pages, en suivant tous les chemins possibles donnés en suivant les liens hypertextes. Il existe pourtant un grand nombre de pages non-indexées, qui constituent ce que l'on nomme le Web invisible ou le Web caché (le Deep Web). Certaines pages, qui n'auraient pas dû être indexées, le sont parfois par erreur.

497 Une version est disponible en ligne sur le Vimeo de Natalie Bookchin, <https://vimeo.com/62458499>, consulté le 30/01/23.

498 Entretien avec Pooja Vyas, par Zoom, en février 2021, <https://sites.google.com/ucsc.edu/femexfilmarchive/filmmaker-index/natalie-bookchin>, dernière consultation le 05/10/21.

499 Une captation vidéo est disponible en ligne sur le Vimeo de Natalie Bookchin, <https://vimeo.com/188841895>, consulté le 30/01/23.

Bookchin crée un rapport intéressant entre l'un des premiers enregistrements sonores — un voyage imaginaire — et ce qui constitue l'une des plus inquiétantes formes d'enregistrement du réel, les images de surveillance. Ces images de vidéosurveillance sont utilisées dans quelques autres films, avant qu'elle ne s'intéresse à une autre nature de vidéo en ligne : les *Vlogs*, vidéos de témoignages sur les réseaux sociaux. Nous abordons cet autre aspect de son travail dans la troisième partie.

## 2.4.2 Paysages visio-sonores : des caméras et des micros sur le monde



Fig. 127 : Jérôme Joy, *Nocinema* (1999-2004)  
© nocinema.org

### ***Le projet Nocinema, capturer les environnements***

Le projet *Nocinema* de l'artiste Jérôme Joy, développé à la toute fin des années 1990 dans le sillage de l'arrivée du streaming, explore l'utilisation de ces nouveaux flux d'images et de sons accessibles par le web. Il s'agit d'un ensemble de projets, articulé au départ autour d'une plateforme web, *nocinema.org*<sup>500</sup>. Lorsque l'on y est connecté, nous voyons des images de paysages naturels ou urbains en plan large et entendons des compositions sonores expérimentales. Il s'agit d'un flux linéaire sur lequel l'utilisateur ne peut agir.

*Nocinema* est, selon les mots de l'artiste, un non-film. En effet, tout comme le *Soft Cinema* de Lev Manovich, il est généré par un programme de montage automatique. À la différence que les images proviennent de webcams disséminées dans le monde.

<sup>500</sup> La page *nocinema.org* ne semble plus directement accessible, mais la documentation se trouve à cette adresse : <https://nocinema.org/htmljs/info/info.html>, consulté le 30/01/23. Tout n'est plus fonctionnel.



Fig. 128 : Jérôme Joy, *Nocinema* (1999-2004)  
Cartographie des sites de webcam  
© nocinema.org

Régulièrement, selon des règles préétablies, le canal de la webcam change automatiquement. Des panoramiques et des zooms automatisés s'opèrent doucement sur les images. Les sons, mixés, proviennent d'une base de données qui est régulièrement nourrie par les compositions d'artistes sonores qui participent au projet. La composition du son ne se fait pas par rapport à des images précises, les artistes ne connaissent pas avant les couplages entre image et son. Ces sons n'obéissent pas à un formatage particulier, si ce n'est celui de l'artiste. Ils peuvent être issus de captations au microphone, ou bien être totalement synthétiques.

Certains mixages fonctionnent sans doute mieux que d'autres selon la nature des sons. Les sons issus de captation microphonique, de phonographies ou *field recordings* amènent leurs « probabilités » avec eux : nous percevons qu'ils sont ancrés dans un environnement qui est le nôtre. Et à la fois, des sons antagonistes ou paradoxaux, synthétiques, instrumentaux ou autres, participent aussi à la construction d'une ambiance, d'une humeur, d'une atmosphère<sup>501</sup>.

*Nocinema* est une sorte de nouvelle forme d'agrégats de paysages filmés, sonorisés, organisée dans un cinéma infini qui se joue sans notre intervention. Entre fiction et do-

501 Jérôme Joy, documentation du projet, <https://nocinema.org/htmljs/info/infonocinema.html>, consulté le 30/01/23.

cumentaire, ces images de webcam, capturées en plan-séquence, sont ici montées par la machine dans une rencontre entre le cinéma et le réseau.

Le projet *Nocinema* existait déjà en 1997 et 1998 sous d'autres formes, avec des fragments de musique générés à partir d'éléments sonores existants, mixés en ligne, ce que Jérôme Joy nommait « musique étendue » ou « hypermusique<sup>502</sup> ». D'autres modules se sont greffés à la plate-forme web *nocinema.org*, visant à explorer d'autres formes mêlant le scénique et les outils de mixage en direct. Certains se mêlent à d'autres dispositifs, comme le *WJs* imaginé par Anne Roquigny, un dispositif de mixage en direct de canaux de sites web<sup>503</sup>. Dans le module le plus récent, *Synema* (2012-2013), un mixage est effectué à partir de sons et d'images capturés grâce à des webcams et des « webmicrophones » (microphones connectés).



Fig. 129 : SLIDERS\_Lab, *Sky Memory Net* (2015)  
© SLIDERS\_Labs

Cette utilisation du son provenant des webcams est relativement rare. Citons les SLIDERS\_Lab<sup>504</sup>, qui dans une dérivation de leur projet *Sky memory Project*, ont mis en forme une version utilisant des flux visuels et sonores en direct. Dans l'œuvre *Sky*

502 *Ibid.*

503 « WJ-S est un logiciel et un dispositif public de performances web, modulable, permettant à des WJ-S (webjockeys, artistes, curateurs, passionnés et mutants du web) de jouer live avec des contenus en ligne. WJ-S est une expérience partagée et visible de la navigation. Les Webjays prennent le contrôle d'un dispositif multiécran et surfent simultanément dans plusieurs fenêtres de navigateurs à la fois. WJ-S est une expérience immersive dans le flux. », <https://www.wj-s.org/about/article/about?lang=fr>, consulté le 30/01/23.

504 Nous avons parlé de leur travail sur les bases de données au chapitre précédent, section 2.3.4.

*Memory Net* (2015)<sup>505</sup>, la base de données n'est donc plus « locale » (c'est-à-dire présente sur une mémoire informatique accessible sans internet), mais elle réside dans un fonds d'images et de sons provenant de webcams connectées à différents lieux dans le monde. L'opérateur navigue mais peut aussi créer en direct des modulations dans les images, ainsi que des variations de formes. Le son joué est récupéré sur les séquences vidéos issues des caméras web.

### ***Que voient et qu'entendent les machines ?***



Fig. 130 : Nicolas Tardy, *L'image contient peut-être*  
© acousticcameras.org

Un autre projet récent, *Acoustic Cameras*<sup>506</sup> (2016), s'inscrit dans la veine de *Nocinema*. Il s'agit d'une plate-forme web, qui propose le visionnage d'une série de pièces audiovisuelles situées<sup>507</sup>. L'image provient d'une webcam qui filme un lieu, en temps réel. Le son est une pièce sonore préparée par un artiste qui a choisi au préalable son lieu. Le son est donc préparé pour l'image, même si celle-ci comporte une part d'imprévisibilité. Un artiste ne peut travailler que sur une seule pièce<sup>508</sup>. Pour chaque pièce sont don-

505 Une version est disponible en ligne sur le site de SLIDERS\_Lab, [http://www.sliderslab.com/pages\\_fr/TRAVAUX/Sky\\_Memory\\_Net.html](http://www.sliderslab.com/pages_fr/TRAVAUX/Sky_Memory_Net.html), consulté le 30/01/23.

506 L'équipe éditoriale du projet *Acoustic Cameras* est constituée de Pierre Belouïn, Christophe Demarthe, Didier Hochart, Nicolas Ledoux et Thierry Weyd.

507 Un peu moins d'une centaine actuellement. Il est possible de filtrer l'index en choisissant de n'afficher que les pièces dont le flux de webcam est disponible.

508 Parmi ces artistes, musiciens expérimentaux, créateurs radiophoniques ou poètes sonores : Inès de Bruyn, Rodolphe Alexis, Thierry Fournier, Christine Webster, Jérôme Poret, Scanner, Pierre Belouïn, Valérie Vivancos, Felix Kubin, Francisco López, Pierre Bastien, Sébastien Roux, Joachim Montessuis,

nés un titre, une description, les coordonnées GPS de latitude et de longitude, et, en temps réel, la date et l'heure.

Le dispositif ainsi créé est la rencontre entre une pièce sonore enregistrée et la captation en direct d'un paysage fluctuant. Selon la saison, selon le temps qu'il fait, selon l'heure de la journée ou de la nuit, la caméra balaie en continu le lieu filmé (bord de mer, ville ou montagne) et écrit en temps réel le film de la musique de l'artiste. La musique, la création sonore, en accaparant le paysage filmé, le revêt aussitôt d'une dimension narrative, dramatique, selon les principes de la musique au cinéma. À la nuance près, qu'ici le film s'écrit en simultané.

Si dans la majorité des « vidéos » la musique prévaut, dans certaines la voix joue un rôle intéressant. Dans *Penser voir* (2018-2022), Thierry Fournier nous donne à voir en plan fixe une plage de la ville de Benicàsim, en Espagne, plus ou moins bondée selon les heures de visionnage. La voix artificielle serait, d'après ce que l'on comprend progressivement, celle de la caméra. Elle tente vainement de décrire ce qui se passe ; mais elle fait face à de nombreux questionnements :

Qu'est-ce que c'est ? Que doit-elle regarder ? Comment reconnaître un paysage ? Qu'est-ce qu'un comportement suspect ? Mais surtout, à quoi sert-elle ? Elle ne comprend rien. Comme si elle était en burn-out, elle se compare aux travailleurs du clic sous-payés qui alimentent les intelligences artificielles, questionnant le sens de son travail, dans un anthropomorphisme qui soulève les enjeux politiques de ces dispositifs<sup>509</sup>.

Toujours dans le cadre de la série *Acoustic Cameras*, mais dans un autre registre, Nicolas Tardy utilise dans sa pièce *L'image contient peut-être* (2021)<sup>510</sup> le flux visuel en plan large d'un ensemble d'immeubles situé dans un quartier d'Amsterdam. L'artiste, de sa voix naturelle, se demande ce qui pourrait se trouver dans l'image. Un même principe régit le texte poétique lu, une énumération de phrases qui débutent chacune par la locution « L'image contient peut-être ». Par exemple, « L'image contient peut-

---

Lionel Marchetti, Charles Pennequin, Rainier Lericolais, RadioMentale, Sylvain Chauveau, P. Nicolas Ledoux, Simon Fisher Turner, Etienne Charry.

509 Thierry Fournier, texte de description de la pièce, <http://www.acousticcameras.org/playlist/thierry-fournier/>, consulté le 30/01/23.

510 Une version est disponible en ligne sur le site d'Acoustic Cameras, <http://www.acousticcameras.org/playlist/nicolas-tardy/>, consulté le 30/01/23.

être... des victimes de la solitude », « L'image contient peut-être... des échanges financiers matérialisés », ou bien encore « L'image contient peut-être... des citoyens perdus dans leurs pensées ».

À elles deux, ces pièces sonores illustrent les problématiques contemporaines liées à la vision des machines, aux algorithmes capables de décrire automatiquement des images, aux problématiques de surveillance et de détection. Nous traiterons ces problématiques par l'angle de la voix et du visage dans la troisième partie.

### ***Les flux sonores du monde***

Depuis de nombreuses années, le laboratoire Locus Sonus<sup>511</sup>, dont Jérôme Joy est cofondateur — développe « un réseau émergent de *streams* (flux audio) environnementaux consacré aux sons singuliers de lieux sélectionnés en Europe et ailleurs<sup>512</sup> ». Ces projets s'inscrivent dans des pratiques de plus en plus courantes entre cartographie sonore et paysage sonore. Il s'agit, comme le prônait Raymond Murray Schafer avec son ouvrage *The Tuning of The World. The Soundscape*<sup>513</sup> (1977), d'attirer l'attention sur les sons environnants et de développer les pratiques de *field recording*, c'est-à-dire l'enregistrement de terrain par l'audio. En direct ou non, ces enregistrements peuvent, comme l'image, témoigner ou s'affirmer comme documents poétiques d'un espace. Ils peuvent être à la fois produits par des artistes, ou bien par des amateurs. Ainsi dans le cas du sonore se sont développés des projets de cartographie sonore en ligne, qui permettent à la fois des possibilités de téléversement (*uploading* ou mise en ligne) de fragments audio situés, ainsi que la possibilité pour des utilisateurs d'explorer les cartes qui les archivent. Noémie Fargier, chercheuse postdoctorale à l'EHESS/CRAL<sup>514</sup>, décrit ainsi ces pratiques qui mêlent la collecte de sons in situ et les *web maps* :

Les cartes sonores en ligne utilisent la technologie des web maps pour associer la prise de son d'un lieu à ses coordonnées géographiques, et ainsi apporter des

511 Locus Sonus est un laboratoire centré sur la pratique des arts sonores. Il a été créé par Jérôme Joy et Peter Sinclair en 2005, au sein de l'école supérieure d'art d'Aix-en-Provence. Un ouvrage relate les expériences qui y ont été — elles continuent de l'être — menées. Jérôme Joy, Peter Sinclair (dir.), *Locus sonus ; 10 ans d'expérimentation en art sonore*, Marseille, Le Mot Et Le Reste, 2015.

512 Description du laboratoire sur le site de l'École d'Aix-en-Provence, <https://esaaix.fr/recherche/locus-sonus>, consulté le 30/01/23.

513 R. Murray Schafer, *Le paysage sonore, op. cit.*

514 Un projet organisé dans le cadre du réseau thématique pluridisciplinaire (RTP) SON:S, lancé en 2019 par le CNRS, coordonné par Karine Le Bail, qui cherche à connecter « les équipes et laboratoires ayant le son et l'écoute pour principal territoire de recherche », <https://sons.cnrs.fr/>, consulté le 30/01/23.

informations sonores sur une localité, dont la carte nous offre un aperçu visuel. Cette pratique de la prise de son en extérieur (dans l'espace naturel ou l'espace public) et cette collecte de sons in situ, communément désignées par le terme de « field recording » [enregistrement de terrain], mais aussi, dans le monde anglophone, par celui de « location sound » [son d'un lieu], sont à la fois liées au développement de l'écologie acoustique initiée par Raymond Murray Schafer et le World Soundscape Project à la fin des années 1960 et à celui de techniques d'enregistrement portatives et création<sup>515</sup>.

Elle ajoute très justement que l'étude de ces dispositifs « tisse des réseaux de réflexion sur la cartographie et l'archive, sur le son, l'écoute et l'audio-vision, sur notre approche sensible des lieux, dans le monde réel et sa représentation, ainsi que sur nos expériences virtuelles du monde, via les nouveaux médias<sup>516</sup> ». Les projets évoqués<sup>517</sup> sont tout à fait passionnants, et impliquent de nombreux acteurs très engagés. On remarquera néanmoins que l'abord esthétique de ces cartes en ligne — ce n'est pas systématiquement le cas — est souvent très classique, au sens où il s'agit d'une carte traditionnelle, plus ou moins retravaillée graphiquement, sur laquelle des points cliquables permettent l'accès aux médias, dans une interface de lecture standard.

---

515 Noémie Fargier, descriptif de la journée d'études « Mappemondes sonores au XXIème siècle. Une autre approche des lieux », Centre de recherches sur les arts et le langage — CRAL, <https://www.ehess.fr/fr/appel-communication/mappemondes-sonores-xxi%C3%A8me-si%C3%A8cle-autre-approche-lieux>, consulté le 30/01/23.

516 *Ibid.*

517 Citons par exemple, en complément de Locus Sonus (<https://locusonus.org/soundmap/051/>), Aporee (<https://aporee.org/maps/>), Écouter le monde (<https://www.ecouterlemonde.net/fr/sonotheque>), Cities and Memory (<https://citiesandmemory.com/>) ou bien encore Nature Soundmap (<https://www.naturesoundmap.com/>), sites consultés le 30/01/23.

### 2.4.3 Fragments géolocalisés. Les *Field\_works* de Masaki Fujihata



Fig. 131 : Masaki Fujihata, *Field\_work Alsace* (2002)  
© antiAtlas des frontières

Formes étendues de la cartographie sonore, d'autres expériences interactives mêlent la carte, le trajet et la base de données, mais aussi l'image et le son. Nous pensons notamment au travail pionnier de Masaki Fujihata, dont nous allons parler dans ce chapitre.

Fujihata est l'un des premiers artistes à avoir utilisé le GPS, dans un projet d'exploration des environs du Mont Fuji. À la suite de cela, il a créé dans les années 2000 une série d'installations interactives dont le contenu, audio-visuel et géolocalisé par des coordonnées GPS, est mis en scène de manière singulière, les *Field\_works*. Voici comment Jean-Louis Boissier, en décrit la teneur :

Il y a d'abord une ligne blanche, complexe, démultipliée, enchevêtrée sur elle-même, celle sur laquelle s'accrochent les images. Il s'agit de la trace du preneur de vues, de Masaki Fujihata, performer équipé d'un sac à dos surmonté d'un récepteur GPS (Global Positioning System), d'une caméra vidéo très légère équi-

pée d'un objectif grand angle à laquelle sont fixés un micro directionnel, un ordinateur de poche et une boussole électronique<sup>518</sup>.

Les *Field\_works*<sup>519</sup> sont une série d'installations audiovisuelles interactives. Chaque installation reflète le trajet parcouru par Fujihata (voire d'autres promeneurs) dans un secteur donné, autour d'une thématique centrale. À différentes étapes du trajet, Fujihata enregistre sous forme audiovisuelle des moments particuliers, des rencontres. Dans la plupart des *Field\_works*, l'image et le son sont couplés d'emblée, par la synchronisation de la saisie d'un caméscope. Ces multiples vidéos sont placées, selon leurs coordonnées, sur une représentation filaire blanche en 3D, qui représente le trajet. Une interface de navigation est proposée à l'utilisateur, décrite ainsi par Douglas Edric Stanley : « un grand bouton rotatif avec lequel l'utilisateur peut se promener en suivant les lignes de la déambulation enregistrée par la machine. Un peu comme un rouleau qu'on déroule dans un pianola, on circule sur les vecteurs de déplacement d'enregistreurs<sup>520</sup> ». Lorsque l'utilisateur navigue, fait tourner et déplacer cet ensemble, les vidéos se situant dans une proximité spatiale viennent occuper les premiers plans — du point de vue l'utilisateur-explorateur — de la cartographie navigable. Or, plus la vidéo est proche du premier plan, plus le volume du son est élevé. Il y a également, dans les versions récentes, un traitement panoramique stéréophonique du son. Il s'agit d'un comportement assez classique, comme dans un jeu vidéo en 3D, dans lequel on avance et dans lequel les sons d'un objet en vue se font de plus en plus présents.

Ici, contrairement à un jeu vidéo d'exploration, les éléments mis en avant sont des vidéos. Une fragmentation spatio-temporelle se produit, et l'on arrive sur une sorte de cluster d'images-sons, sortes de puzzles de temps et d'espaces. Le résultat du mixage automatique du son des différentes vidéos constitue alors le paysage sonore stratifié

---

518 Jean-Louis Boissier, « Masaki Fujihata, "Le linéaire actif". Remarques sur les Field-Works de Masaki Fujihata », [http://www.arpla.fr/canal20/adnm/wp-pdf/Fujihata\\_lineaire\\_actif.pdf](http://www.arpla.fr/canal20/adnm/wp-pdf/Fujihata_lineaire_actif.pdf), consulté le 22/07/2022.

519 Une version du *Field\_work* Alsace (2002) est disponible en ligne sur le Vimeo de Caroline Bernard, <https://vimeo.com/86793927>, consulté le 30/01/23.

520 Douglas Edric Stanley, Article sur son site *Abstract Machine*, <https://abstractmachine.net/en/lexicon/gps>, consulté le 30/01/23.

de zones proches. La spatialisation stéréophonique permettant de mieux distinguer à quel son appartient la vidéo, sorte de relief sonore<sup>521</sup>.

Plusieurs personnes pouvant effectuer des marches avec le système GPS, certaines zones seront plus couvertes que d'autres et proposeront un certain nombre de points de vue.

Plus globalement — et c'est une thématique que nous aborderons dans le troisième chapitre avec le registre audiovisuel des visages et des voix —, la question des relations entre image et son se pose d'autant plus (elle se pose déjà avec des films ou des vidéos linéaires) lorsque l'on agrège, au sein d'installations ou de formes multiécran (comme le *split-screen*), des éléments audiovisuels ayant leur sa propre bande-son. Comment le son individuel et le son collectif résultant sont-ils pensés ?



Fig. 132 : Masaki Fujihata, *Simultaneous Echoes* (2009)  
© YouTube / Masaki Fujihata

Fujihata, qui a pris conscience de la dimension compositionnelle et interactive de ses formes, va tenter, dans *Simultaneous Echoes* (2009)<sup>522</sup>, de préparer de manière plus contrainte la saisie vidéo en allant filmer des musiciens, selon trois règles. La pièce fait

521 Nous supposons qu'avec les nouvelles techniques de spatialisation sonore telle que le son ambisonique, Fujihata aurait pu même jouer encore plus avec la perception sonore, de manière « sphérique ».

522 Une captation vidéo est disponible en ligne sur le YouTube de Masaki Fujihata), <https://www.youtube.com/watch?v=Wi32KBwDzMI>, consulté le 30/01/23.

partie de la série des *Field\_works*, mais propose une autre approche. L'idée qu'il souhaite développer est celle d'une composition multipiste, au sein de laquelle l'utilisateur pourrait « switcher » des éléments, c'est-à-dire substituer certaines pistes à d'autres en naviguant, tout en conservant une sorte d'unité vidéo-musicale générale. Le fait de penser les différentes voix (au sens musical de pistes) invite à penser en termes de composition musicale, mais aussi de combinatoire, puisque différentes vidéos sont susceptibles de se superposer.

Je voulais réaliser des séquences en utilisant des prises de vue enregistrées à des moments différents, mais jouées simultanément, et cela m'a conduit à l'idée d'un morceau de musique. Par exemple, on pourrait enregistrer un contrebassiste dans la rue A, un tambour dans la rue B et un chanteur dans la rue C. Je me demandais si, en les jouant simultanément, on pourrait faire un système de production audio/spatial multipiste<sup>523</sup>.

Le projet s'est fait dans la région de Londonderry en Irlande, avec l'aide du compositeur irlandais Franck Lyons. Comme dans tous les *Field\_works*, un questionnement préalable devient le cœur de la pièce. Dans *Field\_work@Alsace*, il s'agissait des habitants de la région et de leur rapport à la frontière. Ici, le jeu des instruments traditionnels face à leur environnement naturel.

[...] la combinaison de trois types de sons différents, des sons instrumentaux (enregistrements de tambours irlandais et de joueurs de cornemuse), des sons naturels (vent, eau, moutons) et la lecture de partitions musicales, produisent de nouveaux paysages sonores caractérisés par une spatialisation des sons [Nous traduisons]<sup>524</sup>.

Ainsi *Simultaneous Echoes* s'affirme comme une installation-interface tout à fait novatrice, qui entremêle l'idée de composition audiovisuelle interactive, et de navigation dans une base de médias géolocalisés.

523 Masaki Fujihata, « L'avantage technique procure-t-il un avantage expressif ? », in Jean-Marie Dallet (dir.), *Cinéma, interactivité et société*, op. cit., p. 346.

524 « [...]the interplay of three different types of sounds, instrumental sounds of recorded Irish drum and piper players, natural sounds of wind, water and sheep and compositional scores produce novel soundscapes that have the spatialisation of sounds as the main characteristics. », Yvonne Spielmann, « Mapping Scapes of Sound and Vision. Notes on the field work *Simultaneous Echoes*, an audiovisual installation », [http://www.fujihata.jp/s\\_echoes\\_09/ISEA\\_2009\\_Yvonne.pdf](http://www.fujihata.jp/s_echoes_09/ISEA_2009_Yvonne.pdf), consulté le 30/01/23.

## 2.4.4 *Colosseum*. Matérialité du numérique et écologie des média

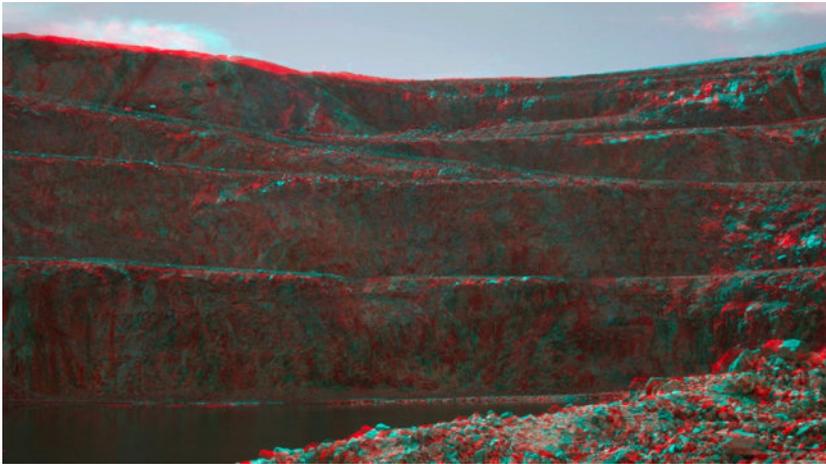


Fig. 133 : Moritz Fehr, *Colosseum* (2015)

Images en 3D

© Moritz Fehr / moritzfehr.de

Dans cette dernière section, nous explorons l'une des faces cachées des paysages technologiques, celle du paysage réel abîmé par les industries du numérique.

Moritz Fehr est un artiste contemporain allemand, qui travaille à la fois le son et l'image dans le cadre d'installations immersives qui mettent en scène des spatialisations singulières. Son installation audio-visuelle *Colosseum*<sup>525</sup> (2015) met en scène une plongée dans une mine abandonnée du désert de Mojave. De celle-ci, abandonnée depuis 1993, étaient extraits de l'or, de l'argent et du cuivre, matériaux nécessaires à la fabrication de réseaux ou de matériel liés aux nouvelles technologies. Actuellement, des minerais encore plus recherchés, les terres rares<sup>526</sup>, sont devenus une matière première du développement de technologies de pointe. La demande étant très forte, certaines mines ont été réactivées, l'une d'entre elles se situant d'ailleurs à proximité de celle de *Colosseum*<sup>527</sup>.

525 En référence au nom de la mine.

526 Ce sont des éléments métalliques malléables appelés lanthanides. Ils permettent la miniaturisation et sont présents, par exemple, dans nos téléphones et nos télévisions. Ces métaux ne sont pas si rares, mais ne se trouvent pas de manière densifiée, mais dispersée, ce qui rend leur extraction plus difficile et consommatrice d'énergie. Ils constituent un enjeu géopolitique majeur, car la Chine concentre en partie leur production.

527 Voir par exemple ce reportage : « Terres rares : des minéraux uniques au cœur d'une bataille mondiale », 02/12/21, Franceinfo, [https://www.francetvinfo.fr/monde/terres-rares-des-mineraux-uniquees-au-coeur-d-une-bataille-mondiale\\_4867259.html](https://www.francetvinfo.fr/monde/terres-rares-des-mineraux-uniquees-au-coeur-d-une-bataille-mondiale_4867259.html), consulté le 30/01/23.

Coulisses symboliques de l'anthropocène, ces gisements dégradent les territoires de manière spectaculaire et inquiétante.

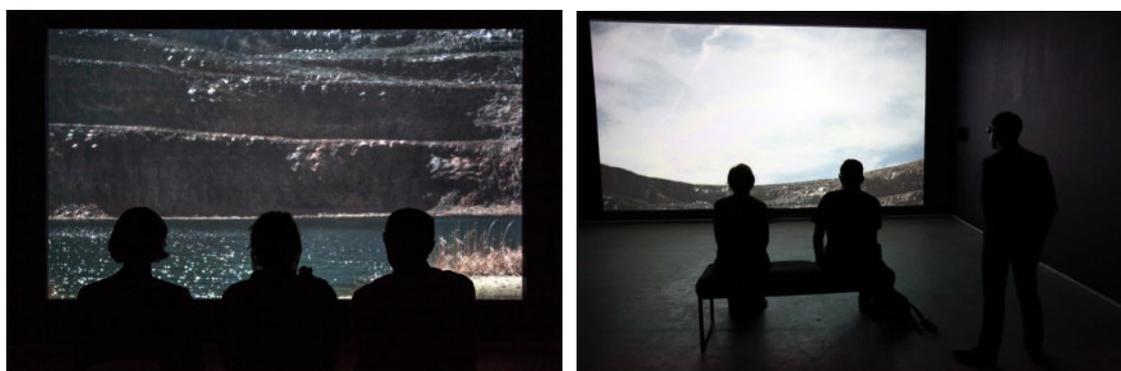


Fig. 134 : Moritz Fehr, *Colosseum* (2015)

Vues de l'installation

© Moritz Fehr / moritzfehr.de

Dans le fond d'un espace sombre, très profond, est diffusée une vidéo numérique en stéréoscopie<sup>528</sup> (des lunettes sont donc nécessaires pour visualiser l'image). Sur les deux côtés latéraux de l'espace sont disposées des enceintes qui diffusent une bande-son spatialisée<sup>529</sup>.

Le film est une descente en caméra subjective, depuis le ciel, jusqu'au fond de la mine de Colosseum, dans lequel stagne un plan d'eau. Le mouvement de caméra dessine une spirale, miroir du relief souterrain dessiné dans le sol par l'évidement de la terre opéré par les bulldozers. La sensation d'immersion dans ce relief est induite par la stéréoscopie, mais aussi par le son, constitué de trois strates de nature différente. La première est a priori une strate ambisonique, c'est-à-dire que le son ambiant de la mine a été capturé avec un ou des micros particuliers, de manière à pouvoir reconstituer le paysage sonore à 360° sur le réseau de haut-parleurs de l'installation. La deuxième strate sonore est constituée de sons produits par l'artiste : des crépitements et des bourdonnements sont obtenus à partir des et des bruits provenant de son ordinateur et de ses disques durs pendant de montage. Les champs électromagnétiques gravitant au sein de l'environnement de travail semblent également être une source de son, à la manière de

528 Que l'on nomme souvent cinéma 3D ou cinéma en relief. Donne à l'image animée une illusion de profondeur.

529 Nous nous basons sur la version installée dans le cadre de l'exposition *Sound Unheard*, Goethe Institut Paris, 2019. Commissaires de l'exposition : Daniele Balit, Katharina Scriba et Anne Zeitz.

l'artiste Christina Kubisch dans ses désormais fameuses *Electric Walks*<sup>530</sup> (2003 —). La nature du processus de fabrication du son, celui du matériel de l'artiste lui-même en train de fabriquer l'œuvre nous évoque également la sculpture de Robert Morris, *Box with the Sound of its Own Making*<sup>531</sup> (1961). La troisième strate sonore achève de créer le trouble. Elle est ce que l'on appelle une gamme de Shepard<sup>532</sup>, c'est-à-dire une composition de sons, ici continue, qui semble monter (au sens : dans l'aigu) indéfiniment, telle une spirale.

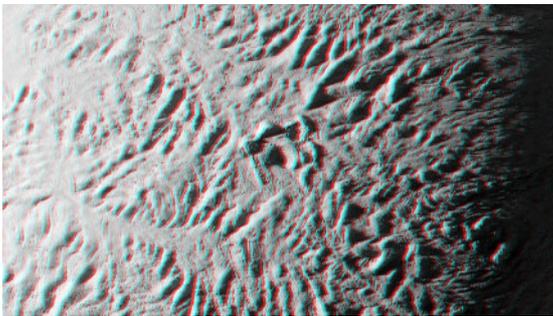


Fig. 135 : Moritz Fehr, *Colosseum* (2015)  
© Moritz Fehr / moritzfehr.de

Les mises en abyme créées par Moritz Fehr sont vertigineuses. D'une part il confronte deux échelles de saisie sonore : celle, extérieure (un paysage sonore) de la mine avec celle, intrinsèque au matériel, de son travail sur le film. D'autre part il lie perceptivement le mouvement de caméra à celui de la gamme de Shepard, en utilisant comme noyau commun la figure de la spirale. L'ensemble donne la sensation de se voir emporté dans une chute sans fin. Fehr fait d'ailleurs référence à l'une des illustrations de *La Divine Comédie* (vers 1310) de Dante, réalisée par Sandro Botticelli : *La Carte de l'En-*

530 Christina Kubisch, artiste allemande pionnière des arts sonores a conçu un casque audio intégrant un procédé de transformation d'ondes électromagnétiques en sons audibles. Pendant des promenades dans l'espace public, à chaque point d'écoute prévu par l'artiste — systèmes de sécurité, guichets bancaires —, les participants peuvent écouter les champs électromagnétiques qui leur étaient inaudibles. L'artiste Max Neuhaus fut le pionnier de ces promenades sonores ou *Soundwalks*, avec sa série *Listen* (à partir de 1966).

531 Il s'agit d'une boîte, assemblage cubique en bois de noyer, posé sur un socle. Grâce à un haut-parleur fixé sur la façade et un magnétophone placé à l'intérieur, elle diffuse « le son de sa propre fabrication », que Morris a enregistré pendant le processus de travail, de trois heures, effectué dans son atelier. On dit que le compositeur expérimental John Cage, de passage à l'atelier de Robert Morris, s'est installé devant la boîte et a écouté la bande entière, comme une composition musicale.

532 L'illusion auditive inventée par Robert Shepard, et reprise dans une version continue par le compositeur Jean-Claude Risset ; notamment dans *Mutations* (1969), bande-son du film éponyme de Lillian Schwartz dont nous avons parlé dans la première partie, section 1.2.5.

*fer*<sup>533</sup> (vers 1480). Celle-ci représente la descente aux enfers de Dante, mené par Virgile, à travers une série d'anneaux de plus en plus étroits. Les voix et plaintes qu'entendent dans le texte les deux personnages pendant cette descente sont en quelque sorte transcrites de manière contemporaine par Fehr. Ce qui est inaudible et invisible directement constitue un danger.

En utilisant des technologies d'image et de sons avancées (le relief sonore et l'image stéréoscopique), Fehr questionne la place grandissante des nouvelles technologies, par un dispositif démonstrateur de la matérialité du numérique. Comme l'analyse Jussi Parika, qui développe une recherche sur l'écologie des média : « Les sols, les sous-sols, les non-sols des infrastructures des média conditionnent le visible et l'invisible. Ce sont des enjeux de relations de pouvoir et de territoire contestés qui mettent en relief le géo au cœur du géopolitique<sup>534</sup> ». Le trou béant laissé par les extractions contraste avec l'idée du numérique comme technologie de l'immatériel.

---

533 Le titre original est *Mappa dell'Infern*.

534 Jussi Parika, « L'Anthroscène », in Jussi Parika, *Trois essais sur l'écologie des media*, Varennes, T&P Publishing, coll. « Iconodules », 2021, p. 39.

## Troisième partie.

### Dialogues et choralités. Les visages-voix de l'ère numérique contemporaine

De toutes les formes artistiques, celle du portrait peint ou photographié est probablement celle qui semble le plus directement s'adresser à nous, dès lors que sa ligne de regard s'accorde avec la nôtre. Le sentiment d'être observé par le sujet représenté, associé à la focalisation induite par ce cadrage particulier — il lui ôte la partie inférieure du corps —, s'est aussi propagé dans l'image animée en étendant son territoire à celui de la voix. C'est ce que pointe Ophélie Naessens, dans son étude sur les portraits parlants (le terme est utilisé en référence aux portraits parlants du précinéma) dans l'art vidéo :

La question du portrait dans ses enjeux contemporains ne peut se penser hors de l'apport du son. Dans ces vidéos, le sujet portraituré parle et nous entendons sa voix. Le récit énoncé, mais aussi sa dimension vocale, les hésitations et les silences, permettent de faire valoir la matérialité de la parole. [...] La voix est synchronisée à l'image, elle est *in*, sa cause est visible dans le champ. Nous attribuons le son à sa source, et ce que nous entendons correspond au mouvement des lèvres que nous voyons simultanément. Notre point de vue et notre point d'écoute coïncident parfaitement. Avec l'image filmique, la voix vient au portrait. La voix est emplie de signifiants et le fait jouer dans le corps du protagoniste. Elle s'incarne, s'inscrit dans les sujets filmés, entre le lieu de leur récit et celui de leur corps. Cette voix, et le timbre qui lui est associé, individualisant le sujet, renforçant la dimension expressive de la représentation<sup>535</sup>.

Avec l'éclosion des réseaux sociaux et de ses témoignages vidéos, ces paroles filmées, aut centrées, ont continué de muter et de proliférer, devenant partie intégrante de notre quotidien. Dans le même temps, la communication avec ces émanations numériques — qu'elles soient des représentations visuelles et sonores capturées depuis le

---

535 Ophélie Naessens, *Portraits parlants dans l'art vidéo. La parole vivante dans les pratiques artistiques des années soixante-dix à nos jours*, Paris, L'Harmattan, 2021, p. 177.

réel (ce que le philosophe Stéphane Vial nomme autruiphonie<sup>536</sup>) ou des simulations numériques (les agents conversationnels) — ouvre de nouveaux questionnements sur les présences à distance, pour reprendre l'expression de Jean-Louis Weissberg.

Aux technologies d'enregistrement s'adjoignent progressivement, on le sait, les technologies de simulation et de téléprésence. Il s'ensuit que de nouvelles générations de spectres font leur apparition. [...] l'ingénierie informatique de simulation a engendré des représentations qui n'obéissent plus à la logique de l'enregistrement, mais à celle de la modélisation. Image d'un présent perpétuellement rafraîchi, elles offrent un espace d'affirmation immédiate et ne restituent plus un ancien présent capturé<sup>537</sup>.

Pour appuyer notre réflexion, nous commenterons plusieurs œuvres installées, dans lesquelles sont donnés à entendre et à voir des voix et des visages. Lorsque ceux-ci seront multiples, nous parlerons de « choralité ». Nous empruntons ce terme au projet collectif de l'Encyclopédie de la parole<sup>538</sup>. Leur site web, qui présente une cartographie de leurs travaux sur l'oralité, propose différentes entrées. L'une d'entre elles, la choralité, est définie ainsi :

Propriété d'une parole dite à plusieurs. Selon que sa forme est plus ou moins unie ou multiple, organisée ou spontanée, la choralité marque soit une subordination à une norme, soit la constitution en temps réel d'un sujet d'énonciation collectif<sup>539</sup>.

Nous étendons ici cette notion à des formes de paroles incarnées visuellement, ou bien à des portraits animés. Nous aborderons également l'interactivité et le dialogue avec des entités audio-écraniques machiniques mettant en jeu la parole et le langage.

---

536 En termes d'interaction interpersonnelle, la manière dont autrui se donne à notre perception (son incarnation dans une présence que l'on expérimente) est dépendante, du point de vue de l'expérience phénoménologique, des conditions techniques liées à sa manifestation. Stéphane Vial la nomme autruiphonie. Entre autres exemples, la conversation téléphonique, qui consiste à « parler sans se voir », se fonde sur les propriétés techno-perceptives de la voix. L'autruiphonie numérique, relative aux manifestations d'autrui à travers les artefacts numériques, est selon lui inédite car elle permet une multitude de communications instantanées, du message textuel jusqu'à la téléconférence audio-visuelle (Stéphane Vial, *L'être et l'écran : comment le numérique change la perception*, Paris, PUF, coll. « Hors collection », 2013, p. 225).

537 Jean-Louis Weissberg, *Présences à distance. Déplacement virtuel et réseaux numériques : pourquoi nous ne croyons plus à la télévision*, Paris, L'Harmattan, 1999, p. 97.

538 L'Encyclopédie de la parole est un collectif qui explore la diversité des formes orales.

539 Site web de l'Encyclopédie de la parole, <https://encyclopediedelap parole.org/fr/entrees/choralites>, consulté le 30/01/23.

## 3.1 Du portrait interactif à l'agent conversationnel

### 3.1.1 Figures du portrait parlant

#### *La photographie de la parole à l'ère du précinéma*

La synchronisation entre image et son a toujours été un axe de recherche parallèle au développement du cinéma muet, comme nous l'avons vu avec Edison. Les premières recherches sont menées par des physiologistes. L'un d'entre eux, Georges Demenÿ, collaborateur d'Étienne-Jules Marey, invente avec Hector Marichelle, professeur à l'Institut National des sourds-muets de Paris, — au cours de ses recherches sur la phonétique expérimentale en 1891 — un appareil nommé phonoscope<sup>540</sup>, qui permet de reproduire, par le biais d'une succession de photogrammes (se résumant à l'époque à des négatifs vagues et sans intensité), une « photographie de la parole ». Le brevet est délivré le 3 mars 1892.



Fig. 136 : Georges Demenÿ, *Photographie de la parole*  
© Gallica

Dans la première version, la caméra chronophotographique (initée par Marey) permettait d'obtenir douze images en celluloïd qui étaient ensuite découpées et collées sur un disque. Ce dernier était placé sur un axe horizontal dans une boîte en bois mu-

540 Giusy Pisano, *Une archéologie du cinéma sonore*, CNRS Éditions/OpenEdition Books, 2004, <https://books.openedition.org/editionscnrs/2753?lang=fr>, consulté le 30/01/23.

nie d'un dispositif optique susceptible d'agrandir les images. Avec l'aide d'une lanterne magique et en tournant une manivelle, on avait alors l'illusion que le visage remuait les lèvres.

Demeny envisage ensuite d'utiliser le phonoscope associé avec un phonographe, de manière à compléter l'illusion, mais ce ne sera pas réalisé. Le 20 avril 1892 s'ouvre l'Exposition internationale de photographie, au cours de laquelle est présentée une version commerciale du phonoscope diffusant une séquence d'images montrant un homme qui prononce la phrase « Je vous aime<sup>541</sup> ».

### ***L'arrivée du cinéma parlant***

Si les intertitres du cinéma muet permettait la dissociation de la parole et du visuel, l'émergence du cinéma parlant à la fin des années 1920, rendu possible par la stabilité d'un système de synchronisation entre image et son, allait ouvrir la porte à ce que Michel Chion nomme un cinéma vococentriste<sup>542</sup>, c'est-à-dire un cinéma qui met systématiquement en avant la voix. De nombreux théoriciens de l'époque voyaient dans ce cinéma émergent, basé sur la parole, une normalisation qui allait contraindre la liberté de rythme dans laquelle s'épanouissent alors bon nombre de films avant-gardistes. La cinéaste et théoricienne Germaine Dulac, parmi d'autres, portait un regard très critique envers le parlant :

Seulement, si je récuse le film parlant, je suis pour le film sonore. Éprise d'harmonie, je sais qu'on peut jouer avec les bruits comme avec les images. [...] Or, ce qu'il y a de terrible et de puéril à la fois, c'est qu'on agit en ce moment à l'égard du film parlant comme on a opéré pour le film muet à ses origines : au lieu d'orienter la découverte vers l'originalité, vers le neuf, on cinématographie uniformément les dialogues. Erreur profonde et qui causera autant de déboires qu'en causa jadis le film muet<sup>543</sup>.

541 Giusy Basile, « Au commencement était le son. Un siècle d'expérimentations avant le parlant », in Nicole Brenez et Christian Lebrat (dir.), *Jeune, Dure Et Pure ! Une Histoire Du Cinéma D'avant-Garde Et Expérimental*, Paris, Cinémathèque Française/Mazzotta, 2001, p. 29.

542 Michel Chion, *La voix au cinéma*, Cahiers du Cinéma, coll. « Cinéma Essais », Paris, 1984, p. 19.

543 Germaine Dulac, « Jouer avec les bruits », *Cinéa-Ciné pour tous*, 1929, in *Germaine Dulac, Écrits sur le cinéma (1919-1937)*, Paris, Paris Expérimental, 1994, p. 129.

Si le scepticisme de Dulac est justifié, le cinéma semblait pourtant voué, si l'on s'en tient aux recherches récentes sur le pré-cinéma, à devenir parlant — Michel Chion parle d'un cinéma muet « en sursis<sup>544</sup> ».

### ***L'illusion du doublage : de la perception multimodale***

Lorsque nous observons une personne en train de parler, et que dans le même temps est diffusé un son qui semble correspondre au mouvement des lèvres de cette personne, alors nous percevons le son comme provenant du même endroit que la source visuelle. Il s'agit d'un cas particulier, pour la parole, de l'ancrage audio-visuel. L'ancrage est défini par Laurent Jullier comme une « relation de causalité établie mentalement par l'auditeur entre un son entendu et une source sonore vue<sup>545</sup> ». Nous intégrons alors les deux signaux en un seul percept.

Ainsi, si nous regardons et écoutons un portrait parlant — et par extension une polyphonie de portraits parlants —, un effet de réalité est à l'œuvre. Nous sommes, en tant qu'être humain regardant et écoutant, sensible à la cohérence de nos perceptions sonore et visuelle. C'est le cas lorsque nous regardons un film doublé, ou bien dans le cas d'un ventriloque et de sa marionnette. C'est pourquoi cette forme d'illusion audiovisuelle a été nommée *Ventriloquism Effect* par les psychologues Ian P. Howard et William B. Templeton en 1966. Dans le cas d'un film, même si le son est spatialisé sur plusieurs enceintes placées à différents endroits de l'espace de diffusion, l'effet reste très fort, le cerveau cherche à rattacher le son à la source visuelle.

Quand les informations visuelles et sonores ne coïncident pas, il peut se produire des illusions perceptuelles. L'un des effets audio-visuels les plus connus est l'effet McGurk. Il a été découvert en 1976 par Harry McGurk et John MacDonald, deux psychologues anglais<sup>546</sup>. Il se produit dans le contexte d'une personne articulant un mot dans un film ou une vidéo, lorsque les informations visuelles et sonores sont rendues contradictoires par l'utilisation d'un montage. Ainsi, si une personne est filmée en train d'articuler le son « ba », mais que le son entendu dans le même temps est « ga »,

---

544 Michel Chion, *La voix au cinéma*, op. cit., p. 24.

545 Laurent Jullier, *Les sons au cinéma et à la télévision : Précis d'analyse de la bande-son*, Paris, Armand Collin, coll. « Cinéma et audiovisuel », 1997, p. 185.

546 Cécile Colin, Monique Radeau, « Les illusions McGurk dans la parole : 25 ans de recherches », 2003, [https://www.persee.fr/doc/psy\\_0003-5033\\_2003\\_num\\_103\\_3\\_29649](https://www.persee.fr/doc/psy_0003-5033_2003_num_103_3_29649), consulté le 30/01/23.

le spectateur peut alors percevoir un mot tel que « da » ou « fa », ou bien une combinaison des deux phonèmes « bga » ou « gba ».



Fig. 137 : Bruce Nauman, *Lip Sync* (1969)  
© Pinault Collection

Ces jeux déstabilisants de son synchronisé ont été explorés par des artistes vidéo comme Bruce Nauman, avec *Lip Sync* (1969)<sup>547</sup>. Cette vidéo présente une performance filmée de l'artiste. En plan inversé vertical, pendant près d'une heure, il essaie de prononcer silencieusement ce qu'il entend au casque — des occurrences répétées de l'expression « Lip Sync ». Le spectateur voit le visage de Nauman, et entend le son du casque.

Le doublage et la postsynchronisation cinématographique sont ainsi des lieux de rencontre sensibles entre la voix et le visage. Comme l'analyse Jacques Aumont :

Le doublage implique une technique, celle du lipping (comment traduire ? Lèvrage ?), soit une coïncidence aussi parfaite que possible entre mouvement des lèvres et phonation, qui signifie à son tour que la voix n'est pas première, mais la parole. Si le désaccord d'une voix et d'un visage, profondément pervers, peut parfois être supporté, au désynchronisme des mouvements des lèvres on ne s'habitue jamais. La bouche parle : cela veut dire que nous voulons qu'elle parle visiblement, que ses yeux soient juges de ce qu'entendent les oreilles. Le lipping a été inventé comme technique du réalisme, mais il est l'instrument de la plus grande irréalité, puisqu'il permet de recoudre une voix à un corps<sup>548</sup>.

547 Une version est disponible en ligne sur le YouTube du GVA IVAM, <https://www.youtube.com/watch?v=hbT9GGIdKOs>, consulté le 30/01/23.

548 Jacques Aumont, *Du visage au cinéma*, Paris, Cahiers du cinéma, coll. « Essais », 1992, p. 53.

La synchronisation entre mouvement des lèvres et voix est donc primordiale dans la production d'un effet de réalité, comme nous le constaterons dans toute cette partie : que ce soient par le biais des capteurs synchronisés de l'ordinateur (la webcam et le micro), la modélisation de visages parlants, ou bien la captation vidéo, la congruence<sup>549</sup> audio-visuelle renforce notre attention.

### ***Le regard-caméra***

Dans un texte consacré au regard-caméra dans le cinéma moderne, Céline Scemama rappelle que si le regard-caméra existe depuis le cinéma muet des frères Lumière, il reste peu usité dans le cinéma classique car il sort le spectateur de son immersion dans l'image. Il se manifeste en revanche de manière plus explicite dans le cinéma moderne d'après-guerre chez des réalisateurs tels que Fritz Lang ou Jean-Luc Godard.

L'inversion du processus est peut-être ce qui fait la grande spécificité de cette forme d'adresse au spectateur à l'ère moderne : ce n'est plus le spectateur qui regarde les images et s'y fonde, mais ce sont les images de cinéma qui regardent le spectateur et l'empêchent ainsi de s'absorber en elles. Le spectateur est alors interpellé en tant que tel, et la fiction lui est en partie refusée<sup>550</sup>.



Fig. 138 : Ingmar Bergman, *Monika* (1953)  
© Telerama

549 Le terme congruence est couramment utilisé en psychologie de la perception. Il explicite un sentiment de bonne coïncidence entre deux événements.

550 Céline Scemama, « You talkin' to me? Sur les regards caméra de la modernité », in Dominique Chateau (dir.), *La direction de spectateurs. Création et réception au cinéma*, Bruxelles, Les Impressions Nouvelles, coll. « Réflexions faites », 2015, p. 226-227.

Chez Godard, c'est un regard-caméra particulier, celui du personnage féminin dans le film *Monika* (1953) d'Ingmar Bergman<sup>551</sup>, qui a inspiré ceux, réguliers, que l'on peut rencontrer dans ses propres films. Sur le point de céder à l'impulsion d'une relation qu'elle avait pourtant réussi à abandonner, par un regard intense, sans paroles, Monika induit une interlocution explicite et contraint le spectateur à se positionner.

### 3.1.2 C'est moi que vous regardez ? Un portrait interactif

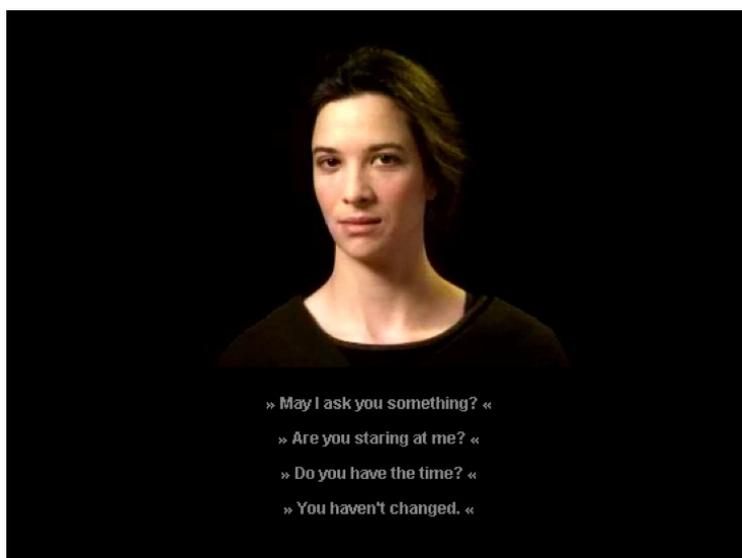


Fig. 139 : Luc Courchesne, *Portrait n° 1* (1990)  
© ZKM

Malgré le regard-caméra, le film de Bergman reste un film au déroulement linéaire. Dans l'une des premières œuvres reposant sur un dispositif numérique, l'artiste Luc Courchesne va proposer la potentialité d'influer sur le cours du film, dans le contexte de ce que l'on peut nommer un « cinéma interactif<sup>552</sup> ». *Portrait n° 1* (1990)<sup>553</sup> met en scène une jeune femme trentenaire (nous apprendrons qu'elle se nomme Marie), cadrée en « plan rapproché poitrine<sup>554</sup> », isolée sur un fond noir. Sans intervention du vi-

551 Jean-Luc Godard, « Monika », in Jean-Luc Godard, *Godard par Godard. Les Années Cahiers*, Paris, Flammarion, 1989, p. 146-148.

552 On considère souvent le film *One Man and his World* (1967) de Raduz Cincera comme le premier film interactif (non numérique). Un animateur sur la scène demandait au public présent dans la salle, à différents moments, de choisir entre deux bifurcations possibles du scénario.

553 Une capture vidéo est disponible en ligne sur le Vimeo de Luc Courchesne, <https://vimeo.com/5827424>, consulté le 30/01/23.

554 Ou P.R.P. Dans le vocabulaire du cinéma, un plan dont le cadre cible le haut du corps, depuis juste au-dessous des épaules jusqu'au haut du visage. Ce type de plan est souvent utilisé pour les dialogues. Avec les communications du type visio-conférence, ce format devient un standard. Il laisse deviner un

siteur, l'écran présente une image fixe de Marie portant son regard hors champ, vers la gauche du cadre, comme en attente.



Fig. 140 : Luc Courchesnes, *Portrait n° 1* (1990) © Fondation Langlois



Fig. 141 : Luc Courchesnes, *Portrait n° 1* (1990) © Media Art Net

Sous l'image se trouvent des textes cliquables, chacun correspondant à des bifurcations<sup>555</sup> possibles de la conversation. Après avoir cliqué sur le premier texte « Excusez-moi<sup>556</sup> », le buste du personnage de Marie pivote doucement et silencieusement (le bruit de fond de la captation vidéo, qui se déclenche au passage de l'image fixe à l'image animée nous laisse néanmoins supposer que la vidéo sera sonore) afin de diriger son regard vers le visiteur. Le regard de Marie évoque à ce moment celui de Monika.

Les quatre textes (dont trois questions) qui s'affichent alors définissent des orientations explicitement différentes : « Puis-je vous poser une question ? », « C'est moi que vous regardez ? », « Auriez-vous l'heure ? », « Vous n'avez pas changé. ». Effectivement, selon les chemins parcourus, la conversation est orientée vers des pensées phi-

espace en arrière-plan, mais n'en montre pas trop. C'est un cadrage ni trop rapproché, ni trop éloigné, qui correspond à une réalité. Nous pourrions tout à fait nous filmer en plan éloigné avec un micro et une rallonge, le son resterait le même ; mais il ne correspondrait pas, dans le rapport image-son avec cette reproduction sociale d'un face à face, au même niveau.

555 La bifurcation est l'une des figures centrales de l'interactivité selon Jean-Marie Dallet. Elle est « le lieu où une voie de communication se dédouble » (Jean-Marie Dallet, « Figures de l'interactivité », in *Revue Anomalie* n°3, « Interfaces », Clermont-Ferrand, Un, Deux... Quatre Éditions, 2003, p. 37).

556 C'est la seule option sur ce premier écran, les deux autres textes n'ayant pour objet que d'activer un sous-titrage.

losophiques ou des rapports de séduction, mais peut tout aussi bien s'écourter brusquement.

L'installation se compose principalement d'un socle-cabine sur lequel est posé un moniteur d'ordinateur, ainsi que des haut-parleurs discrets. Tout est fait pour que le haut du corps du visiteur soit en face de celui de Marie et corresponde à sa posture. Dans certaines versions, une plaque de verre réfléchissante permettait à la fois de mettre en avant les actions textuelles possibles tout en créant une strate d'intimité plus forte entre le visiteur et Marie. Les premières versions étaient stockées sur un vidéodisque, l'un des tout premiers supports permettant une navigation interactive dans des médias stockés. Une version en ligne, reprogrammée, est disponible en ligne<sup>557</sup> ; à l'heure des avatars numériques, l'expérience reste encore très troublante.

Pour Miguel Almiron, artiste et théoricien, Marie dévoile dans son discours sa véritable identité : « seulement un portrait fixé dans la durée du temps de l'enregistrement qui a été joué par un sujet de chair et d'os<sup>558</sup> ». Malgré cette artificialité donnée, il pointe l'expérience qui se crée par cet échange :

Marie a seulement un passé, mais elle se présente virtuellement dans un continuum espace-temps de rencontre à l'aide des appareils technologiques. C'est l'instrument technologique ici qui s'efface de notre regard pour commencer ce jeu ; ainsi un dialogue avec un personnage virtuel s'établit pour proposer une possible existence conjointe<sup>559</sup>.

Cette existence conjointe définit, pour Luc Courchesne, ce qu'il nomme un art expérimental. Dans celui-ci le spectateur joue un rôle fondamental : il monte les séquences d'images-sons.

Mes installations vidéo interactives ont souhaité interpeller le spectateur. C'est d'abord ce dernier que j'ai toujours voulu mettre en scène. Dans cette perspective, la conversation demeure à mon avis une situation toute naturelle pour entraîner le spectateur dans un échange virtuel. Bien que les questions et les ré-

557 La pièce est accessible sur le site de la Fondation Langlois, <https://www.fondation-langlois.org/Portrait-no-1/>, consulté le 30/01/23.

558 Miguel Almiron, *À la Recherche du visage perdu dans l'art numérique*, Villeneuve-d'Ascq, Presses universitaires du Septentrion, coll. « Arts du spectacle — Images et sons », 2022, p. 98.

559 *Ibid.*

ponses soient prédéterminées, les choix, eux, demeurent libres. Le spectateur doit pouvoir faire son propre cinéma<sup>560</sup>.

Avec *Portrait n° 1*, Courchesne investit de manière pionnière les interfaces conversationnelles entre personne et ordinateur, en programmant une architecture de bifurcations potentielles. Dans la prochaine section, nous explorons la manière dont certaines de ces interfaces se sont peu à peu autonomisées. D'abord par le texte, puis par anthropomorphisation avec les voix et les visages artificiels.

### 3.1.3 Les *chatterbots*. Premiers dialogues avec l'ordinateur

Comme le spécifie Luc Courchesne, dans ses pièces telles que *Portrait n° 1*, les questions et les réponses sont prédéterminées. En termes d'interactivité, l'utilisateur est contraint à effectuer un choix parmi quelques propositions. D'autres méthodologies de conversation dématérialisée, plus génératives, et utilisant l'apprentissage automatique, ont également émergé à partir des années 1960.

#### *Naissance de l'intelligence artificielle*

En 1956, la conférence « The Darmouth Summer Research Project on Artificial Intelligence », à laquelle participait entre autres le mathématicien et ingénieur Claude Shannon (à l'origine du concept de théorie de l'information en 1948<sup>561</sup>), marque officiellement la naissance des explorations sur l'intelligence artificielle (I.A.), dans la lignée des premières recherches d'Alan Turing<sup>562</sup>. Cette approche se démarque de la cybernétique<sup>563</sup>, au sens où il ne s'agit pas d'arriver à modéliser des phénomènes neurologiques, mais d'utiliser directement la computation de l'ordinateur existant comme producteur de cognition. Cela peut se comparer au moment où l'on passe des auto-

560 Luc Courchesne and Bernard Perron, « T'es où [Luc Courchesne] ? », *Intermédialités/Intermediality* n° 9, p. 157–175, 2007, <https://doi.org/10.7202/1005535ar>, consulté le 30/01/23.

561 Claude Shannon, « A Mathematical Theory of Communication », in *The Bell System Technical Journal* n° 27, New York, AT&T, 1948.

562 Jean-Claude Heudin, *Les Créatures artificielles : Des automates aux mondes virtuels*, Paris, Odile Jacob, 2008, p. 139.

563 Concept défini notamment par le mathématicien, chercheur et théoricien Norbert Wiener, dans ses ouvrages *Cybernetics* (1948) et *The Human Use Of Human Beings* (1950). Il y définit une science d'étude des mécanismes de contrôle et de communication dans des systèmes complexes, naturels ou artificiels, en particulier les relations entre l'homme et la machine (« Norbert Wiener on cybernetics », in Jasia Reichardt [dir.], *Cybernetic Serendipity. The computer and the arts*, catalogue d'exposition, Studio International Fondation, New York/Washington, 1968–2018, p. 9).

mates qui essaient de reproduire la voix par la reconstitution d'un conduit vocal à la synthèse vocale, qui est une simulation générative.

Au-delà de la puissance de calcul rendue disponible, ces premières interactions ont aussi fait émerger la question du dialogue avec la machine, et de ses interfaces. À la fin des années 1980 naissent les I.R.C. (Internet Relay Chat), l'ancêtre du *chat* (bavardage en français) qui permettent des discussions entre utilisateurs à partir d'un petit logiciel client installé sur l'ordinateur. Le besoin de modération sur les salons de discussion a très vite fait émerger des robots modérateurs, qui simulent une présence et répondent à des questions<sup>564</sup>.

### **La psychologie d'ELIZA**

Le premier *chatterbot*<sup>565</sup> (contraction de *chatter robot*) semble être *ELIZA*, créé vers 1966. Il a été programmé en SNOBOL par Joseph Weizenbaum, chercheur au MIT (Massachusetts Institute of Technology) au sein du laboratoire dédié à l'intelligence artificielle. Il s'agissait pour Weizenbaum d'arriver à reconstituer des échanges inspirés des techniques du psychologue Carl Rogers, qui se fondent sur la répétition ou la reformulation des problèmes abordés par le patient. Le programme, relativement simple, analyse les phrases saisies par l'utilisateur, et les reformule sous forme de questions. Le principe de l'écholalie<sup>566</sup> est au centre de la programmation. S'il ne comprend pas, il affiche ironiquement « Je comprends ». Jean-Claude Heudin, dans son ouvrage *Les créatures artificielles* en cite un extrait :

ELIZA : Pourquoi dites-vous cela maintenant ?

LE PATIENT : Je ne sais pas.

ELIZA : Réfléchissez. Approfondissez votre réponse.

LE PATIENT : C'est à cause de mon père.

ELIZA : Dites-m'en plus à propos de votre père<sup>567</sup>.

<sup>564</sup> Jean-Claude Heudin, *op. cit.*, p. 208.

<sup>565</sup> Emmanuel Cyriaque, « Écritures, lectures : nouvelles formes », in Frédéric Migayrou (dir.), *Coder le monde*, *op. cit.*, p. 103.

<sup>566</sup> « Trouble du langage qui consiste à répéter de manière systématique les derniers mots entendus », définition du Larousse.

<sup>567</sup> Jean-Claude Heudin, *op. cit.*, p. 142.

Il n'y a pas de réelle intelligence derrière ce chatbot, c'est un leurre. Mais beaucoup d'utilisateurs semblaient se laisser prendre au jeu. Weizenbaum, dans une démarche presque artistique souhaitait montrer la superficialité du dialogue d'un humain avec la machine. Chez le sujet, une anthropomorphisation est à l'œuvre. Le professeur en médias informatiques Noah Wardrip-Fruin, spécialiste des programmes informatiques et des jeux vidéo, nomme « Effet Eliza » cette « capacité d'un dispositif informatique à suggérer à un auditoire sa complexité illusoire<sup>568</sup> ».

En 1972, Kenneth Colby crée cette fois-ci un agent conversationnel susceptible d'imiter le comportement d'un sujet schizophrénique. En 1973, ELIZA et ce patient artificiel se sont rencontrés par l'initiative de Vint Cerf, pionnier du web<sup>569</sup>, marquant le début d'une ère d'exploration du dialogue autonome entre machines.

Un agent conversationnel comme *Alice*<sup>570</sup>, projet débuté en 1995, est toujours développé, et est diffusé sous une forme libre et exploitable permettant à tous d'utiliser son intelligence artificielle dans le cadre de projets allant de l'interface client dans des applications web commerciales à des projets artistiques.

### 3.1.4 Détacher la voix du corps : paroles artificielles

Parmi les évolutions majeures des bots, la capacité à « parler » a été la première étape de l'anthropomorphisation des bots, grâce à la synthèse vocale. D'un point de vue communicationnel, la parole se présente, selon Philippe Breton, sous l'angle d'une double articulation avec l'intériorité (le dialogue intérieur, vers soi) et avec le monde extérieur (la communication par le langage). Elle serait donc à la fois le lieu d'une distance à soi, mais aussi celui d'une distance au monde.

La parole est donc à la fois ce qui supporte et qui crée cette distance au monde. Elle est l'instance qui nous permet de réduire cette distance, mais en même

---

568 Marc Jahjah, « Digital et Book Studies : les formes de la matérialité numérique », 2015, <https://marcjahjah.net/198-digital-book-studies-22-formes-de-materialite-numerique>, consulté le 30/01/23.

569 Xavier Aimé, « Intelligence artificielle et psychiatrie : noces d'or entre Eliza et Parry », in *L'information psychiatrique* n° 93, 2017, p. 51-56, <https://www-cairn-info.ezpaarse.univ-paris1.fr/revue-l-information-psychiatrique-2017-1-page-51.htm>, consulté le 30/01/23.

570 ALICE (Artificial Linguistic Internet Computer Entity). Le réalisateur Spike Jonze cite ce système comme inspiration pour son film *Her* (2014). Dans celui-ci, le héros, Theodore, tombe amoureux de Samantha, un système artificiel incarné par une voix de synthèse.

temps de la fonder. La parole nous éloigne et nous rapproche dans le même mouvement<sup>571</sup>.

Avec les techniques de gravure sur cire qui aboutirent à l'invention du phonographe en 1877 s'est opérée la première fixation du son. Axée à sa naissance sur la voix parlée, cette machine parlante ou machine à fantômes a permis à la voix de devenir dans le même temps éternelle, et manipulable. Cette même période, qui voit aussi naître le téléphone, puis une vingtaine d'années plus tard la radio, est qualifiée négativement par le compositeur et théoricien Raymond Murray Schafer de « schizophonique » (en grec *schizo* signifie fendre, *phônê* désigne la voix), au sens où l'on sépare le son de sa source originelle<sup>572</sup> ; il devient transmissible et reproductible, et n'est plus associé à un espace ou à un temps précis. C'est pour lui le début d'une ère qui va laisser se développer une sorte de pollution sonore, dans laquelle « les sons naturels sont de plus en plus remplacés par des sons artificiels, et où les signaux qui ponctuent la vie moderne ne sont plus que des substituts fabriqués par les machines<sup>573</sup> ». Cette description du « son artificiel », appliquée à la voix, est décrite de manière analogue par Bruno Bossis, musicologue, dans son ouvrage *La voix et la machine*. Pour lui, lorsque des sons de paroles entendus par un auditeur proviennent d'une machine, ceux-ci doivent être considérés comme artificiels<sup>574</sup>, quand bien même ils sont basés sur un enregistrement de voix réelle ; la voix naturelle, fixée, se déshumanise et devient phénomène. Bossis cite à ce propos Pierre Schaeffer :

L'homme qui enregistre sa propre voix et qui se met à l'écouter comme celle d'un autre, se trouve dans une situation que l'enregistrement seul permet et qui présente un caractère profondément inhumain<sup>575</sup>.

L'histoire de la vocalité artificielle débute toutefois bien avant celle du phonographe, avec les premières machines parlantes qui cherchaient à simuler la voix humaine grâce à des systèmes mécaniques reproduisant de manière plus ou moins précise l'appareil

571 Philippe Breton, *L'éloge de la parole*, Paris, La Découverte, 2007, p. 52.

572 Raymond Murray Schafer, *Le Paysage sonore. Le monde comme musique*, op. cit., p. 141.

573 *Ibid.*, p. 143.

574 Bruno Bossis, *La Voix et la machine. La vocalité artificielle dans la musique contemporaine*, Rennes, Presses Universitaires de Rennes, coll. « *Æsthetica* », 2005, p. 26.

575 Bruno Bossis, op. cit., p. 248.

phonatoire humain. En 1780, l'abbé Mical présentait rue du Temple à Paris deux têtes parlantes. Utilisant les techniques mécaniques de reproduction de la voix (déjà évoquées dans les sections précédentes) basées sur des glottes artificielles, et contrôlables par des claviers, les têtes pouvaient dialoguer ou répondre au public<sup>576</sup>. Dès ses premières, la voix artificielle est déjà intimement liée au dialogue.

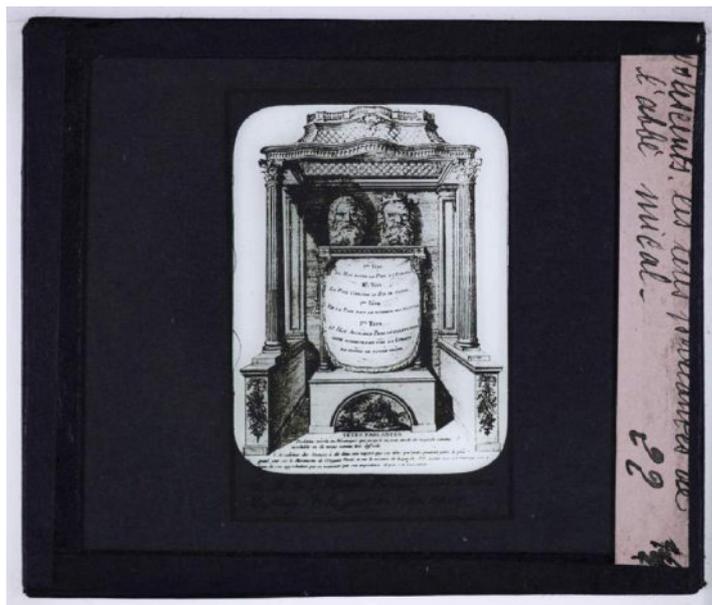


Fig. 142 : Les Jeux des adolescents (les automates). Les têtes parlantes de l'abbé Mical.  
© Collections du Musée national de l'Éducation

Autre exemple, la machine du baron Wolfgang Von Kempelen pouvait en 1835<sup>577</sup> prononcer de manière approximative des phonèmes basiques. Elle s'appuyait sur un soufflet et une boîte à vent dont les effets étaient modulés grâce à des sifflets, des valves et des leviers. De nombreuses inventions (plus fidèle au corps humain, avec des claviers de contrôle) ont vu le jour, mais ont très vite montré leurs limites. Un changement de paradigme, pointé par le chercheur en culture sonore Jonathan Sterne, survient : celui du passage de la bouche à l'oreille<sup>578</sup>, grâce aux progrès effectués dans la recherche sur l'acoustique. L'un des pionniers dans ce domaine est Hermann von Helmholtz. Son synthétiseur, vers 1850, est capable de produire des sons vocaliques, non plus en es-

576 Jean-Claude Heudin, *op. cit.*, p. 66.

577 Reconstitution en 1967 par un chercheur du CNRS Jean-Sylvain Liénard.

578 Jonathan Sterne, *Une histoire de la modernité sonore*, Paris, La Découverte/Cité de la musique, coll. « Cultures sonores », 2015, p. 116.

sayant de reproduire la cause d'un son, mais en travaillant sur ses effets. Entre autres innovations, le *Vocoder* inventé par Homer Dudley et présenté en 1940 permet l'analyse et la synthèse de la parole. Il sera utilisé durant la Seconde Guerre mondiale par l'armée américaine pour communiquer.

Quelques années après, les premières expériences de synthèse numérique sont menées aux Bell Telephone Laboratories durant la période 1957-1961, notamment avec l'ingénieur John Pierce et le chercheur Max Mathews. Si la recherche débute sur le codage vocal, dont Shannon avait théorisé la possibilité<sup>579</sup>, il faut attendre 1962 pour que les premières avancées importantes dans la synthèse vocale numérique se concrétise, l'une de ses manifestations célèbres étant la fameuse chanson *Bicycle Built for Two*<sup>580</sup>. Le superordinateur *Hal*, dans le film de 2001, *l'Odyssée de l'espace* (1968) de Stanley Kubrick, chantonne cette chanson dans ses derniers moments, lorsque sa mémoire est vidée<sup>581</sup>. L'accompagnement de la chanson est réalisé en sons de synthèse par Max Mathews, mais la voix artificielle est produite par des ingénieurs, qui ont créé *Blodi*, un logiciel indépendant de *Music III* (le logiciel de synthèse musicale utilisé aux Bell Labs à l'époque). *Blodi* était capable de produire une parole artificielle à partir d'entrées phonétiques.

Des musiques intégrant la vocalité artificielle vont apparaître dans les années 70, celles du compositeur Charles Dodge dans ses *Speech Songs* par exemple. Il est l'un des premiers compositeurs à utiliser les voix de synthèse informatiques dans sa musique. Tout à fait conscient de l'aspect robotique des voix, il joue avec ces sujets dans les textes mêmes mis en jeu dans beaucoup de ses pièces. *The Story of Our Lives* (1974) est un dialogue numérique entre un homme et une femme, qui lisent ensemble, sur un canapé, un livre relatant l'histoire de leur vie<sup>582</sup>. Au fil de la pièce, ils réalisent le vide que constitue cette vie et essaient de s'en départir. Les voix des deux protagonistes

---

579 Olivier Baudouin, *Pionniers de la musique numérique*, Sampzon, Delatour, 2012, p. 58.

580 Parfois nommé *Daisy Bell* ou *Daisy*, *Daisy* (Olivier Baudouin, *op. cit.*, p. 75).

581 Olivier Baudouin rappelle toutefois que cette version a pris la version synthétique comme modèle, mais qu'il s'agit en réalité d'un ralentissement de la voix d'un chant a cappella d'un l'acteur (Olivier Baudouin, *op. cit.*, p. 76).

582 Notes by Charles Dodge, in *Synthesized Speech Music by Charles Dodge*, Original liner notes from CRI LP jacket, <https://nwr-site-liner-notes.s3.amazonaws.com/nwcr1348.pdf>, consulté le 30/01/23.

chantent, se dédoublent, se multiplient, glissent de l'une à l'autre. Ce chant à deux voix est régulièrement interrompu par une autre voix, neutre et irréaliste, celle du livre. Peu à peu, elle vient dominer celle du couple.

Dans un contexte d'artificialité, la question de la présence est centrale. L'archéologue des médias Wolfgang Ernst en parle ainsi :

C'est évident quand on analyse la voix, et qu'on se rapporte à l'invention du phonographe. Le phonographe a démontré que les voix étaient faites d'ondes sonores périodiques, qui pouvaient être artificiellement recréées. Il nous est alors devenu possible de penser la voix humaine comme un genre de composition technique. C'est une autre façon de comprendre les sens. La culture numérique, bien qu'elle semble de prime abord complètement multimédia et multisensorielle (les gens utilisent l'ordinateur de manière visuelle, auditive et tactile), est paradoxalement complètement non-sensuelle — c'est là le véritable pouvoir de l'ordinateur. L'ordinateur n'est pas accessible au travers des sens. Les opérations mathématiques qui le fondent sont entièrement non-sensuelles. Elles sont de l'ordre du sublime, pour employer le langage esthétique du dix-huitième siècle. Il est possible de sentir une présence, qui pourtant ne peut être appréhendée par les sens<sup>583</sup>.

Au-delà de cette présence recherchée, lorsque l'agent conversationnel est incarné par un visage parlant, un système technologique étendu et transversal doit être mis en place. Il faut reconnaître et analyser les entrées de l'utilisateur (textuelles ou vocales), puis en retour synthétiser une parole, effectuer la synchronisation labiale (*lip sync*) et adapter les expressions faciales.

Tous ces processus, lorsqu'ils savent se rendre invisibles, donnent lieu à des expériences de dialogue inédites. C'est le cas de *DiNA*, un agent conversationnel incarné imaginé par l'artiste Lynn Hershman Leeson.

---

583 Élodie A. Roy, « Retours critiques sur l'archéologie des médias : un entretien avec Wolfgang Ernst », in *Cahier Louis-Lumière* n° 10, « Archéologie de l'audiovisuel », Paris, ENS Louis-Lumière, 2016, p. 20.

### 3.1.5 DiNA, un agent conversationnel de Lynn Hershman Leeson

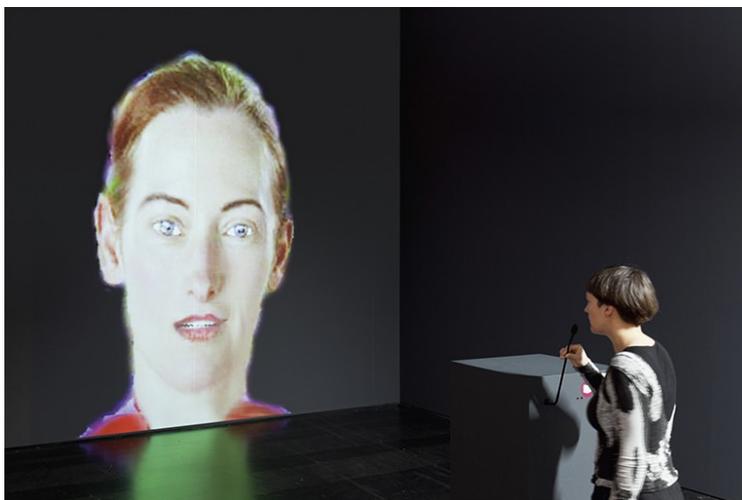


Fig. 143 : Lynn Hershman Leeson, *DiNA* (2004-2006)  
© Lynn Hershman Leeson / lynnherhman.com

Les questions d'identité, de persona et d'exposition de soi sont au cœur du travail de l'artiste américaine Lynn Hershman Leeson, depuis ses premières pièces plastiques — des figurines sonores<sup>584</sup> — jusqu'à l'exploration du biogénétique dans ses dernières pièces. Après le développement au début des années 70 d'un personnage fictif — Roberta Breitmore —, elle explore en pionnière<sup>585</sup> la relation interactive permise par le vidéodisque<sup>586</sup> avec *Lorna* (1979-1983), le personnage d'une femme quarantenaire agoraphobe et solitaire, recluse dans son appartement<sup>587</sup>. Le jeu voyeuriste du regard est également pour elle, à cette période, un axe de réflexion important. Elle l'explore à

584 Au début des années 70, au University Art Museum de Berkeley, l'une de ses premières expositions de poupées (qui intégraient un magnétophone diffusant des respirations et des rires) a été subitement arrêtée car le sonore n'avait, selon le commissaire, pas sa place au musée (Jean Gagnon, « Lynn Hershman », Fondation Langlois, 2006, <https://www.fondation-langlois.org/html/f/page.php?NumPage=168>, consulté le 30/01/23).

585 Il s'agirait de l'un des tous premiers vidéodisques interactifs artistiques, avec ceux de Luc Courchesne (*Portrait n° 1*) et celui de Jean-Louis Boissier (*Le Bus*).

586 Les premiers vidéodisques, pas encore numériques, permettaient un accès direct interactif à différentes séquences vidéos analogiques stockées sur un support disque.

587 Les seuls moyens de communication de Lorna vers l'extérieur sont le téléphone et la télévision, qui sont pour elle anxigènes (notamment la sonnerie, les publicités ou les actualités). Après avoir pris place dans une installation comportant quelques éléments de mobilier et de décoration qui reconstituent l'appartement, le visiteur peut naviguer dans l'interface visuelle diffusée sur une télévision grâce à une télécommande. Les objets affichés dans l'interface (portefeuille, télévision, téléphone, aquarium, miroir...) sont accessibles par les touches numérotées de la télécommande ; à chaque objet correspond une séquence vidéo et sonore qui se joue et dévoile des indices. Le fil de ses interactions (il y a 36 séquences-chapitres) va lui permettre de découvrir la vie de *Lorna*, d'accéder à son intimité, et peut mener à différentes issues : le suicide, le statu quo (l'enfermement volontaire de Lorna se prolonge) ou bien la destruction de la télévision par le personnage, acte symbolique de rébellion.

travers des œuvres en réseau telles que *CyberRoberta* (1996), une poupée à l'effigie de son persona Roberta Breitmore qui intègre une webcam et dont le regard est contrôlé par les internautes, ou bien avec *Une chambre à soi* (1990-1993), dispositif proche d'un peep show (mais elle fait également référence au kinéscope d'Edison), dans lequel on fait pivoter une œillère mobile qui provoque des changements de séquence d'écran sur le mur au fond d'un décor de poupée. À partir de 1986, Hershman Leeson entame ses *Electronic Diaries*, vidéos dans lesquelles elle se filme et se confie à la caméra.



Fig. 144 : Lynn Hershman Leeson  
*Lorna* (1979-1983)  
© ZKM

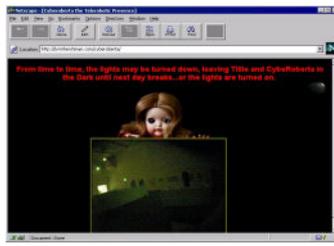


Fig. 145 : Lynn Hershman Leeson  
*CyberRoberta* (1996)  
© anthology.rhizome.org



Fig. 146 : Lynn Hershman Leeson,  
*Electronic Diaries* (1986 —)  
© New Museum

Dans la continuité de ces travaux, le numérique va lui permettre de mettre à profit l'intelligence artificielle (par apprentissage) dans *Agent Ruby* (1998-2002), puis *DiNA* (2004-2006), qui matérialise un certain nombre de problématiques. *Agent Ruby* était un agent conversationnel en ligne, qui permettait aux internautes de se livrer.



Fig. 147 : Lynn Hershman Leeson, *Agent Ruby* (1998-2002) © SFMOMA



Fig. 148 : Différentes expressions d'*Agent Ruby*  
© Fondation Daniel Langlois

Les échanges textuels ont été ensuite rejoués par l'actrice Tilda Swinton (qui joue plusieurs clones de la même entité) dans le film *Teknolust* (2002). Lorsque l'on demande à

Agent Ruby<sup>588</sup> qui est son créateur, elle répond Rosetta Stone (« I was created by Rosetta Stone. »), une scientifique célèbre (« She is a famous computer scientist. »). Rosetta Stone est en effet la généticienne qui crée l'intelligence artificielle dans le film. Selon le dictionnaire des arts médiatiques, la notion d'agent a été introduite vers 1950 par John McCarthy du MIT, et a été réutilisée par Brenda Laurel au début des années 1990 dans sa théorie de l'interaction entre les humains et les ordinateurs. Voici une partie de cette définition :

Entité de nature logicielle (ou robot logiciel) effectuant de façon relativement autonome certaines actions dans le cadre d'un programme multimédia ou de réalité virtuelle [...] Les agents ont souvent un caractère anthropomorphique. Ils sont exprimés visuellement par un ensemble de traits correspondant aux attentes et à la culture des utilisateurs et permettant d'inférer leur mode d'action ou de rôle<sup>589</sup>.



Fig. 149 : Lynn Hershman Leeson, *DiNA* (2004-2006)  
© bitforms gallery

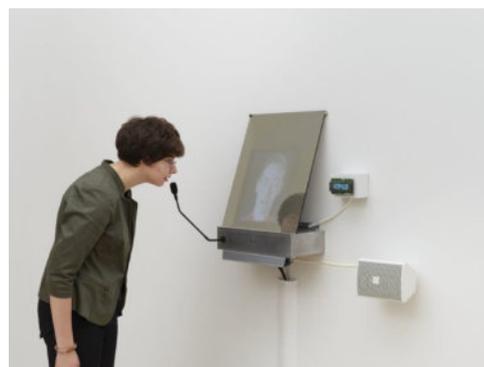


Fig. 150 : Lynn Hershman Leeson, *DiNA* (2004-2006)  
© SEQUITUR / Galerie Bridget Donahue

*DiNA*<sup>590</sup> est en quelque sorte une évolution personnifiée de *Agent Ruby*, elle est capable d'échanger avec un visiteur grâce à un dispositif d'*input* (détection de la voix), et d'*output* (génération de voix artificielle synchronisée sur une image animée de visage). *DiNA* était même à l'origine conçue pour candidater à l'élection d'un téléprésident<sup>591</sup>.

588 Une version en ligne, reprogrammée et fonctionnelle est disponible à cette adresse :

<http://agentruby.sfmoma.org/>, consulté le 30/01/23.

589 Louise Poissant (dir.), *Dictionnaire des arts médiatiques*, Québec, Presses de l'Université du Québec, 1997, p. 8.

590 Une captation est disponible en ligne sur le YouTube de la bitforms gallery, à 9 min 42 s,

<https://www.youtube.com/watch?v=8Yggbr0cPXE>, consulté le 30/01/23.

591 Lynn Hershman, « DiNA – The People's Choice », Media Art Net, 2004,

<http://www.medienkunstnetz.de/source-text/150/>, consulté le 30/01/23.

Il était prévu qu'elle puisse être installée dans une cabine isolée afin que l'on puisse voter pour elle. Elle est au courant des informations du Web, qui sont régulièrement mises à jour dans sa mémoire. Enfin, son visage animé reflète les mots prononcés en temps réel. Le logiciel *Pulse 3D veepers* (qui n'existe plus aujourd'hui) est celui qui anime le visage (en particulier les lèvres) et les fait coïncider avec la voix artificielle. L'échange, quant à lui, est géré par *Alicebot*<sup>592</sup>, un système programmable de *chat*. L'écran montrant *DiNA* est placé sur un miroir sans tain, de manière assez similaire au *Portrait n° 1* de Luc Courchesne dont nous avons parlé dans une section précédente.

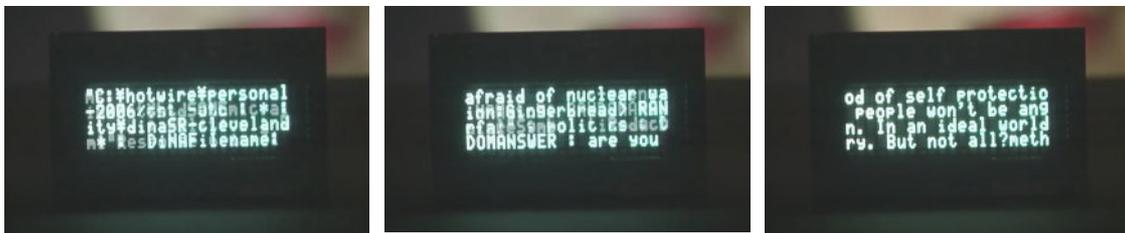


Fig. 151 : Lynn Hershman Leeson, *DiNA* (2004-2006)

L'écran qui diffuse les « pensées » de *DiNA*, à trois moments différents

© Whitney Museum / YouTube

Un petit écran, sur le côté de l'installation, permet de visualiser la réflexion de *DiNA*, comment s'opèrent les associations. Il est possible de voir ce qu'elle interprète, mal ou non, ou bien encore ce qu'elle s'apprête à dire.

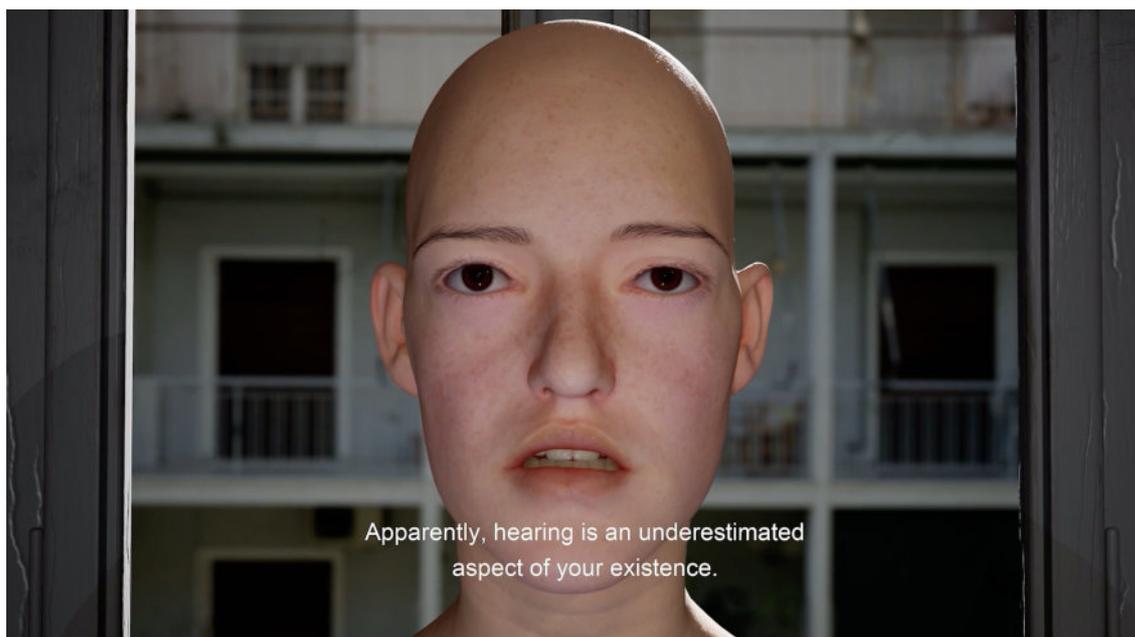
Dans une conférence donnée avec Trevor Paglen, l'artiste Hito Steyerl donne l'exemple de deux chatbots de Facebook auxquels on a appris à négocier de manière commerciale, autour d'un objet<sup>593</sup>. Au bout d'un certain temps, les chatbots ont mis en place une sorte de métalangage. Ce que Steyerl nomme « interlangue » est ainsi un langage interne qui se met en œuvre, dans un fonctionnement qui reste opaque. Observer les images de pensée de *DiNA* relève d'une expérience similaire.

592 Nous avons évoqué ce système dans l'une des précédentes sections, la 3.1.3.

593 Conférence d'Hito Steyerl et de Trevor Paglen, *Les images comme « Social Projectors »*, Centre Pompidou, 10 novembre 2018, <https://www.centrepompidou.fr/fr/ressources/media/10fUnFF>, consulté le 30/01/23.

Nous pouvons ainsi nous questionner sur que les machines voient, entendent ou pensent. Avec sa vidéo *Not allowed for algorithmic audiences*, l'artiste Kyriaki Goni nous donne cette opportunité.

### 3.1.6 Sur écoute : confessions d'un assistant artificiel



« Apparemment, l'écoute est un aspect sous-estimé de votre existence. »

Fig. 152 : Kyriaki Goni, *Not allowed for algorithmic audiences* (2021)  
© Kyriaki Goni, copie privée

Pendant sept jours, avant son arrêt définitif, un assistant personnel virtuel<sup>594</sup> se livre et explique les raisons réelles de son existence. Un monologue vidéo d'une dizaine de minutes est diffusé chaque jour, l'ensemble constituant le développement par étapes de la réflexion de l'assistant. Dans cette œuvre, *Not allowed for algorithmic audiences* (2021)<sup>595</sup>, l'artiste grecque Kyriaki Goni<sup>596</sup> met en scène un être artificiel non genré, chauve, dont les mouvements des lèvres sont synchronisés avec une voix — qui semble féminine —, enregistrée<sup>597</sup>. Des indications textuelles d'heure et de température, ainsi

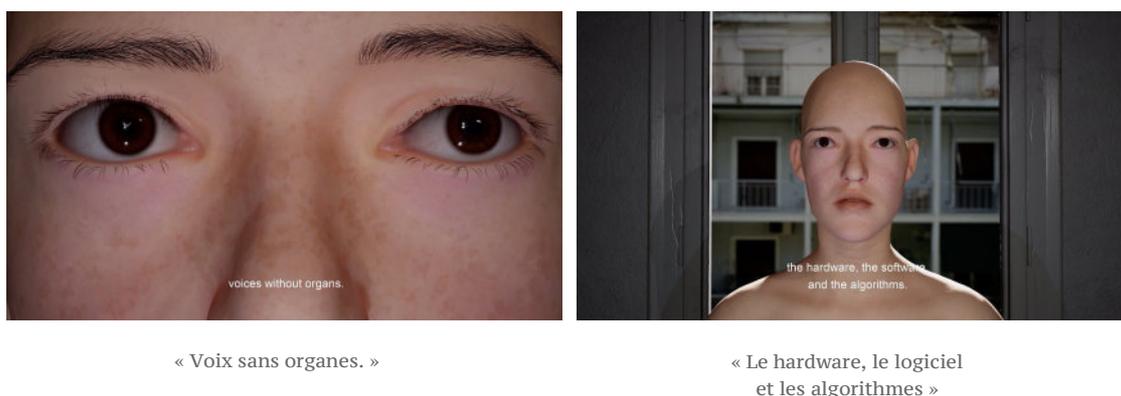
594 En anglais Intelligent Personal Assistant (IPA)

595 Un extrait est disponible en ligne sur le Vimeo de Kyriaki Goni, <https://vimeo.com/604759242>, consulté le 30/01/23.

596 L'artiste nous a mis à disposition une captation complète des monologues.

597 Le modèle, le même pour la voix et le visage, est une actrice grecque, Sofia Kokkali.

qu'un arrière-plan (la façade et les balcons d'un immeuble visible depuis ce qui semble être l'appartement de l'agent) viennent infirmer un semblant de réalité. L'ensemble, dans sa mise en scène, s'apparente à la vidéo de quelqu'un qui se sait menacé. La voix, monocorde, est posée, et les mouvements faciaux relativement réalistes ; les clignements de paupière, notamment, créent une empathie assez immédiate. De temps à autre, l'assistant se tait. Trois types de plan sont alternés : plan serré sur les yeux, gros plan, plan poitrine.



« Voix sans organes. »

« Le hardware, le logiciel et les algorithmes »

Fig. 153 : Kyriaki Goni, *Not allowed for algorithmic audiences*, 2021  
© Kyriaki Goni, copie privée

L'histoire est racontée à la manière d'un conte, ou d'une histoire de ses origines. L'agent nous dit que nous pouvons l'appeler *Voice* (Voix). Elle évoque tout d'abord les premières conceptions du robot dans la mythologie à travers la figure de Talos<sup>598</sup>, ou bien encore des femmes artificielles, faites en or, qui pouvaient parler et marcher. L'un des thèmes du monologue devient de plus en plus clair lorsqu'elle évoque Athanasius Kircher. Celui-ci avait inventé la « statua citofonica » (statue parlante). Grâce à un système de long tuyau en spirale, il était possible, pour celui qui possédait la statue, de récupérer des voix dans un espace public pour les retransmettre dans l'espace privé où se situait la statue. Le public écouté ne se doutait pas qu'il était écouté.

Le philosophe et musicologue Peter Szendy cite aussi l'exemple de Kircher dans son ouvrage *Sur écoute* (2007). Il ajoute que Jeremy Bentham, philosophe anglais qui imagine au XVIII<sup>e</sup> siècle le *Panopticon*<sup>599</sup>, avait envisagé d'y intégrer un pendant « pana-

598 Géant de bronze qui, dans certains mythes, aurait été un automate forgé par Héphestos, le dieu grec du feu.

599 Un système de surveillance panoptique fondée sur une architecture carcérale circulaire permettant aux gardiens de surveiller les prisonniers sans que ceux-ci ne sachent à quels moments ils sont obser-

coustique » ; à savoir un système de tuyaux permettant d'écouter ce qui se passe dans les cellules<sup>600</sup>. En l'état, le système présentait toutefois le défaut majeur de donner la possibilité à l'écouté d'entendre le surveillant.

Au fil de l'histoire sont progressivement tracés des ponts entre les systèmes d'écoute et les systèmes de vocalité artificielle. Puis le spectateur-auditeur est invité à réfléchir à la manière dont elle (l'agent) a été créée, au fait que ces systèmes deviennent de plus en plus petits, invisibles ; aux systèmes de commande vocale tels que Siri. Derrière l'artificialité se cache une matérialité, des composants qui utilisent des matériaux rares, non renouvelables, qu'il est nécessaire d'extraire dans différents pays. Les biais de genre et raciaux induits par les systèmes de détection automatique sont également dénoncés :

Des études ont montré qu'aucune erreur ne se produit quand il s'agit d'identifier le genre d'un homme avec une peau claire. Alors que plus d'erreurs se produisent quand il s'agit d'une femme ou d'une personne à la peau plus foncée. Selon une étude récente, les systèmes de reconnaissance vocale ont plus de mal à reconnaître la voix de personnes noires. (Selon une étude récente).

L'agent admet ensuite être entraînée pour recueillir et reconnaître les voix, par leur tonalité, leur timbre, leur prosodie. Toutes les caractéristiques des sons que nous produisons sont analysées en permanence, ajoutées à des bases de données. La voix devient l'interaction la plus utilisée, selon elle, dans les rapports entre humain et machine. Les intelligences artificielles sont capables de détecter des sons inhabituels, comme le bruit d'une vitre brisée. Kyriaki Goni semble ici faire allusion à une installation vidéo de l'artiste Hito Steyerl, *The City of Broken Windows* (2018). Celle-ci est constituée de deux projections, de forme documentaire, face à face<sup>601</sup>. L'une d'entre elles nous montre des séries d'essais de « brisage » de vitre par les employés d'une so-

---

vés. Michel Foucault analyse notamment en 1975 cette structure de contrôle dans son ouvrage *Surveiller et punir. Naissance de la prison* (Sophie Limare, *Surveiller et sourire : les artistes visuels et le regard numérique*, Montréal, Les Presses de l'Université de Montréal, 2015, p. 12-13).

600 Peter Szendy, *Sur écoute. Esthétique de l'espionnage*, Paris, Minuit, coll. « Paradoxe », 2007, p. 36.

601 Une exposition consacrée à l'artiste, *Hito Steyerl. I will survive. Espaces physiques et virtuels*, a eu lieu au Centre Pompidou en 2021. Dans la version de *The City of Broken Windows* présentée, les deux projections sont suffisamment éloignées l'une de l'autre afin que le visiteur ne comprenne pas d'emblée que les deux projections sont liées.

ciété. On comprend à la vision de l'autre projection qu'il s'agit d'une entreprise de vidéosurveillance. En testant, analysant et classant les différents bruits de vitre brisée, ils peuvent détecter plus efficacement les effractions. L'objet sonore, au sens de Pierre Schaeffer, sort du champ musical pour être mis au service de l'industrie de la surveillance.

Pour conclure, l'agent nous donne des conseils pour éviter l'indexation systématique et déstabiliser les machines : produire des sons différents, faire varier nos énonciations, inventer de nouveaux langages.

Dans ces nouveaux modèles fondés sur l'apprentissage profond (*Deep Learning*), nos voix et nos visages sont constamment détectés, capturés, analysés et classifiés. Par les machines cela va de soi, mais également par l'humain, qui agit en souterrain comme un « dresseur d'intelligence artificielle<sup>602</sup> », avec tous les biais et les effets de bord qui en peuvent en découler.

## 3.2 Visages et voix diagrammatiques

### 3.2.1 Anatomies de la voix comme commande

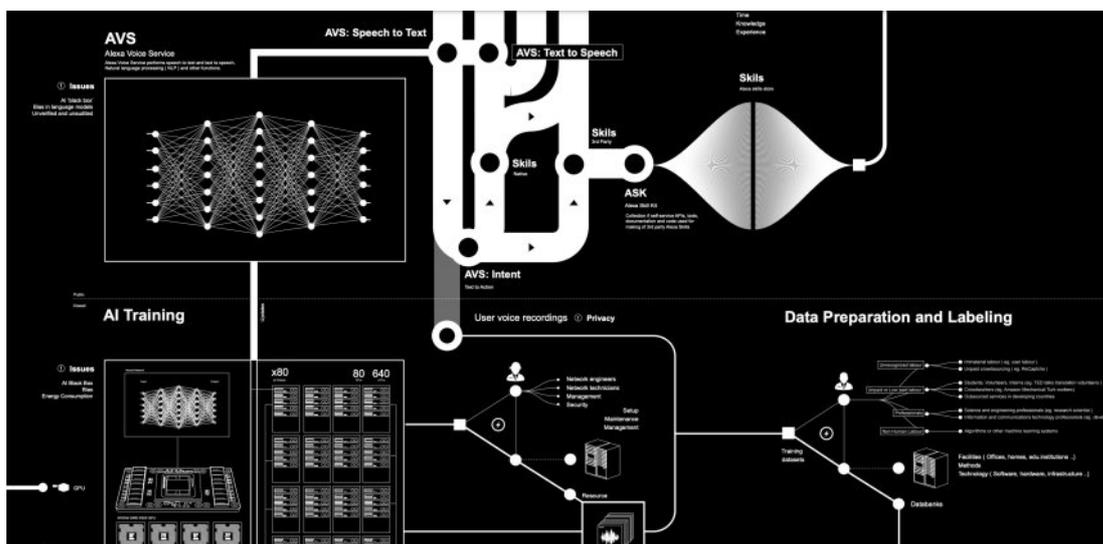


Fig. 154 : Kate Crawford et Vladan Joler, *Anatomy of an AI System* (2018), détail.  
© anatomyof.ai

602 Antonio Casilli, « L'avènement du dresseur d'intelligences artificielles », *Le Monde*, octobre 2011, [https://www.lemonde.fr/economie/article/2020/02/06/qui-travaillera-demain-les-robots-ou-les-humains\\_6028592\\_3234.html](https://www.lemonde.fr/economie/article/2020/02/06/qui-travaillera-demain-les-robots-ou-les-humains_6028592_3234.html), consulté le 30/01/23.

À la fin de la partie précédente, nous expliquions comment Kyriaki Goni, à travers l'histoire de son agent *Voice*, nous mettait en garde sur l'utilisation des sons que nous donnons aux grandes compagnies. C'est aussi ce que Christian Marclay, avec ses *Sound Stories*<sup>603</sup>, cherche à faire en pointant le côté sonore des *Vlogs*. En effet, si le grand public semble un minimum sensibilisé aux problématiques de l'image partagée<sup>604</sup>, la question du sonore reste couramment associée à la radiophonie et aux communications téléphoniques. Une prise de conscience s'est toutefois opérée avec le développement des assistants vocaux.

À la fin des années 2010, des enquêtes ont ciblé les pratiques de traitement de données liées à Alexa, l'assistant vocal d'Amazon. Celles-ci ont montré qu'une partie des enregistrements étaient envoyés à des sous-traitants, retranscrits et annotés par des « petites mains » humaines (des microtravailleurs) dans différents pays, afin d'aider l'intelligence artificielle à se mieux se former<sup>605</sup>. Seulement ces enregistrements sont se sont avérés parfois très sensibles, certains provenaient même de conversations privées dont l'enregistrement s'était malencontreusement déclenché. À la suite de ces découvertes, il s'est avéré que cette situation était similaire chez tous les GAFA<sup>606</sup>, même si certains affirmaient découper et rendre anonymes les échantillons sonores. Depuis, la situation semble s'être améliorée, des politiques de consentement ont émergé, mais les écoutes continuent car l'intervention humaine est nécessaire à l'apprentissage des machines. La Quadrature du net<sup>607</sup> avait publié le témoignage de l'une de ces travailleuses chargées de transcrire les conversations, accompagné d'un texte d'Antonio Casilli, professeur en sociologie et chercheur en humanités numériques<sup>608</sup>. Ce dernier y expose l'opacité de ces systèmes et les risques qui découlent des opérations humaines souterraines.

---

603 Voir dans la deuxième partie, section 2.3.5.

604 C'est-à-dire les mutations de l'image dans l'environnement numérique du Web et des réseaux sociaux. Ce concept est notamment développé dans l'ouvrage d'André Gunthert, *L'Image partagée. La photographie numérique*, Paris, Textuel, 2015.

605 « Les assistants vocaux d'Apple, Amazon et Google vont (un peu) moins vous écouter », Le Monde, août 2019, [https://www.lemonde.fr/pixels/article/2019/08/05/assistants-vocaux-apple-amazon-et-google-levent-le-pied-sur-les-ecoutes-humaines\\_5496728\\_4408996.html](https://www.lemonde.fr/pixels/article/2019/08/05/assistants-vocaux-apple-amazon-et-google-levent-le-pied-sur-les-ecoutes-humaines_5496728_4408996.html), consulté le 30/01/23.

606 Acronyme désignant les principaux acteurs du numérique : Google, Apple, Facebook, Amazon.

607 Association activiste fondée en 2008 autour des questions liées au numérique, notamment celles de la censure et de la surveillance.

608 Voir aussi son ouvrage, consacré à ce sujet : Antonio A. Casilli, *En attendant les robots, Enquête sur le travail du clic*, Paris, Seuil, coll. « La Couleur des idées », 2019.

Les logiciels à activation vocale tels Cortana, Siri ou Alexa sont des agents conversationnels qui possèdent une forte composante de travail non artificiel. Cette implication humaine introduit des risques sociétaux spécifiques. La confidentialité des données personnelles utilisées pour entraîner les solutions intelligentes est à risque. Ces IA présupposent le transfert de quantités importantes de données à caractère personnel et existent dans une zone grise légale et éthique. Dans la mesure où les usagers des services numériques ignorent la présence d'êtres humains dans les coulisses de l'IA, ils sous-estiment les risques qui pèsent sur leur vie privée. Il est urgent de répertorier les atteintes à la *privacy* et à la confidentialité associées à cette forme de « digital labor », afin d'en estimer la portée pour informer, sensibiliser, et mieux protéger les personnes les plus exposées<sup>609</sup>.

Pour lui, ces systèmes dits d'« intelligence artificielle » vont mener au développement de travail précaire et moins protégé<sup>610</sup>.

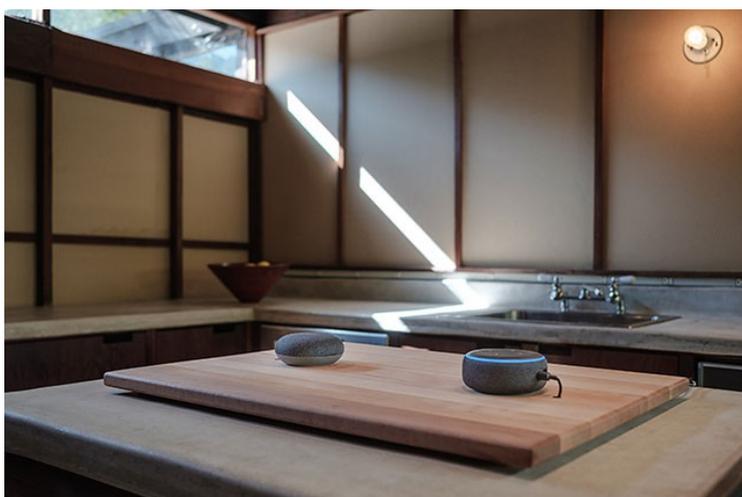


Fig. 155 : Jeff Guess, *Seamlessness* (2019)  
© MAK Center for Art and Architecture

Pour éclairer le fonctionnement opaque de ces systèmes, différents chemins sont empruntés par les théoriciens et les artistes. En 2018, les deux professeurs en nouveaux médias Kate Crawford et Vladan Joler ont dressé une cartographie de l'écosystème de

609 Antonio Casilli, « Les humains derrière Cortana », dossier « Derrière les assistants vocaux, des humains vous entendent », *La Quadrature du Net*, mai 2018, [https://www.laquadrature.net/2018/05/18/temoin\\_cortana/](https://www.laquadrature.net/2018/05/18/temoin_cortana/), consulté le 30/01/23.

610 Comme il l'explique, ces retranscriptions sont déjà majoritairement effectuées par des travailleurs recrutés sur des plate-formes de « micro-travail », qui proposent du travail à la tâche, sorte de mutation numérique du taylorisme.

l'assistant vocal Alexa<sup>611</sup>. *Anatomy of an AI System*<sup>612</sup> nous présente ainsi l'ensemble des ressources humaines, technologiques et énergétiques nécessaires au fonctionnement d'une telle intelligence artificielle<sup>613</sup>. L'artiste Jeff Guess a quant à lui, dans *Seamlessness* (2019)<sup>614</sup>, fait dialoguer deux enceintes connectées concurrentes, l'Amazon *Echo* et la Google *Home*. La discussion, en temps réel mais scénarisée, portait sur leurs conditions respectives d'existence. Comme l'analyse le poète et créateur sonore David Christoffel, les enceintes connectées font partie d'un écosystème très fermé<sup>615</sup>. Il est difficile à cette heure de les configurer autrement que par les personnalisations limitées offertes par des applications dédiées (pour *Alexa*, elles sont appelées les *Skills*). Son travail va alors consister à explorer de manière poétique le rapport entre l'humain et la machine. Faire dialoguer des prisonniers avec une intelligence artificielle, dont le rôle est d'encourager le locuteur, mais qui ne comprend pas les émotions complexes ou ambiguës. Moins « sérieusement », faire dialoguer deux étudiants : l'un pose des questions farfelues telles que « Pourquoi les boutiques ouvertes 24 h/24 et 7j/7 ont-elles des serrures ? » ; l'autre joue le rôle de l'assistant artificiel qui doit répondre à ces questions. Le trop-plein d'empathie forcée vis-à-vis du locuteur dans le premier cas, ou bien l'absence de second degré dans le deuxième cas surgissent immédiatement à l'écoute des créations radiophoniques issues de ces expériences.

Malgré la mise en lumière des arcanes systémiques des assistants numériques par les artistes et les chercheurs, nos échanges vocaux deviennent peu à peu des marchandises. La sociologue Shoshana Zuboff, dans son ouvrage *L'âge du capitalisme de surveillance*, consacre de nombreuses pages à cette question de la voix « qui s'insinue

---

611 Une série de douze articles sur Alexa et les assistants vocaux a été publiée par le média web Les Jours : Sophian Fanen, « Alexa, l'espionne qui m'aimait. Le web devient vocal, mais qui nous parle ? », <https://lesjours.fr/obsessions/alexa-amazon-espionne/>, consulté le 30/01/23.

612 La cartographie est disponible à cette adresse : <https://anatomyof.ai/>, consulté le 30/01/23.

613 Les systèmes *Text-to-Speech* (synthèse vocale à partir du texte) et *Speech-to-Text* (reconnaissance de la parole) y sont bien entendu centraux.

614 Une captation est disponible en ligne sur le site de Jeff Guess, <https://www.guess.fr/works/seamlessness>, consulté le 30/01/23.

615 David Christoffel, « La voix des GAFA », intervention dans le cadre de la journée d'étude « Vox Machines », organisée sous la responsabilité scientifique de Julien Drochon et Anthony Masure par le Pôle Nouveaux Médias de l'École supérieure d'art des Pyrénées (ESA, Pau-Tarbes) et le Master Design Transdisciplinaire, Cultures et Territoires, Université Toulouse Jean-Jaurès, 10 décembre 2018, <https://www.canal-u.tv/chaines/universite-toulouse-jean-jaures/vox-machines/la-voix-des-gafa-david-christoffel>, consulté le 30/01/23.

entre votre existence et les nouveaux marchés qui en font leur cible<sup>616</sup> », et qui devient, à travers la « conversation », un objet de convoitise pour les GAFAs.

Google s'est rapproché d'autres entreprises technologiques désireuses d'imposer la « conversation » comme canal d'échange entre les humains et le dispositif. À terme, cette obsession de la voix sera peut-être dépassée ou rattrapée par d'autres : ainsi, il suffira d'émettre une pensée ou d'agiter un doigt pour déclencher une action. Pour l'heure, c'est la parole qui constitue le Graal [...]. La voix qui vise à la domination, la Voix Unique, deviendra la géante des pipelines de surplus comportemental ; elle en tirera un avantage concurrentiel potentiellement insurmontable, tant sera grande sa capacité à piéger et à s'accaparer la partie dominante de l'expérience humaine. La « conversation » est la seule pour l'heure à promettre une domination de l'approvisionnement en matière première ; les gains de la Voix Unique devraient être astronomiques<sup>617</sup>.

Selon elle, il s'opère une chasse à la parole, qui est pour ces grandes compagnies le nouveau Graal. Elles doivent s'approvisionner et constituer des réserves de flux de conversations, pour aider le marketing et la publicité à mieux analyser les comportements et prédire les réflexes d'achat des utilisateurs.

Plus les machines extraient de ces surplus de paroles des aperçus structurels, plus le flux de commerce qui en émane est important. Ce qui signifie que la valeur de ce que vous dites ne peut être matérialisée sans machines qui apprennent de la manière dont vous dites les choses. Cette forme de surplus provient de la structure de vos énoncés : vocabulaire, prononciation, intonation, rythme, inflexions, régionalismes<sup>618</sup>.

La voix est devenue une donnée comme une autre, monétisable, revendue et analysée pour encore mieux cibler les comportements. Elle qui relevait encore récemment d'une sorte de tradition radiophonique rassurante, s'inscrit progressivement, avec le numérique, dans les mêmes environnements de circulation, de surveillance et de détection que ceux de son alter ego visuel, le visage.

---

616 Shoshana Zuboff, *L'âge du capitalisme de surveillance*, Honfleur, Zulma, coll. « Essais », 2022, p. 353.

617 *Ibid.*, p. 352.

618 *Ibid.*, p. 355.

### 3.2.2 Détections biométriques de la déviance : typage des visages et reconnaissance faciale



Fig. 156 : Alphonse Bertillon, Planche de visages de criminels  
© France Info

Comme l'explique la théoricienne et critique d'art Marion Zilio dans son ouvrage *Faceworld*, avec l'arrivée de la photographie, et depuis Alphonse Bertillon et sa base de visages de détenus<sup>619</sup>, les processus de typage de profil n'ont cessé de se développer<sup>620</sup>. Il s'est agi de toujours plus discrétiser les éléments du visage pour en classer les multiples combinaisons. La numérisation de l'image a en ce sens catalysé ce courant s'évertuant à transformer le visage en objet technique analysable : « La traduction du visage en données numériques et son indexation dans des répertoires initièrent une vision diagrammatique, tout à la fois descriptive et statistique<sup>621</sup> ». Avec l'avènement des bases de données numériques, des procédés comme la reconnaissance faciale sont devenus un lieu stratégique pour les entreprises de marketing.

L'un des écueils de ces approches de typage et de classification est le biais. James Bridle, théoricien et artiste, souligne que la technologie n'est pas neutre, qu'elle encode les préjugés des humains qui la développent<sup>622</sup>. Ce qui est présenté comme un travail automatique de la machine, comme pour la commande vocale, est en réalité

619 À la fin du XIX<sup>e</sup> siècle, le criminologue Alphonse Bertillon invente une méthode de classification anthropométrique des criminels à partir de la photographie de leur visage.

620 Marion Zilio, *Faceworld. Le visage au XXI<sup>e</sup> siècle*, Paris, PUF, coll. « Perspectives critiques », 2018, p. 34.

621 *Ibid.*, p. 35.

622 James Bridle, *Un nouvel âge de ténèbres. La technologie et la fin du futur*, Paris, Allia, 2022, p. 169.

toujours porteur d'une idéologie sous-jacente. Quand Zilio évoque Bertillon, Bridle cite le psychométricien Francis Galton, l'inventeur dans les années 1880 du portrait « composite<sup>623</sup> ». Il s'agit d'une technique d'exposition multiple, fondée sur la fusion de plusieurs prises de vues de visages. Galton induit l'idée qu'il pourrait y avoir, par exemple, des visages types de malades, ou de criminels. Pour autant, Bridle n'accuse pas les grandes compagnies qui utilisent ces nouvelles technologies de reconnaissance de créer volontairement des biais, il pointe le fait que les machines d'apprentissage utilisent des données orientées.

Il est plus probable que ces machines révèlent les inégalités systémiques toujours présentes au sein de la main-d'œuvre des entreprises technologiques, où ceux qui développent sont majoritairement blancs. [...] Elles révèlent également, comme jamais auparavant, les préjugés historiques profondément encodés dans nos ensembles de données, qui structurent notre savoir et nos décisions actuelles<sup>624</sup>.

Comme l'analyse l'artiste et géographe Trevor Paglen, les images sont de plus en plus créées par des machines, pour des machines, et les humains de moins en moins impliqués dans la boucle<sup>625</sup>. Son travail s'inscrit dans le sillage du cinéaste Harun Farocki qui a exploré dans ses films, à travers le concept d'« image opératoire », les nouveaux rapports qui s'établissent entre l'œil et la machine<sup>626</sup>. Paglen constate que les villes, les aéroports, les centres commerciaux intègrent des dispositifs de surveillance susceptibles de capter et d'interpréter des visages, des mouvements, voire des émotions. Dans plusieurs de ses œuvres, Paglen examine les travers et les biais de ces dispositifs de reconnaissance faciale.

---

623 *Ibid.*, p. 167.

624 *Ibid.*, p. 171.

625 Conférence d'Hito Steyerl et de Trevor Paglen, *Les images comme « Social Projectors »*, *op.cit.*

626 Notamment avec sa série de trois installations, *Eye/Machine I-III* (2000-2003), dans lesquelles il met en scène des archives d'images militaires ; par exemple des images prises par des missiles autoguidés pendant la guerre du Golfe.



Fig. 157 : Trevor Paglen, *Machine-Readable Hito* (2017)  
© Trevor Paglen / paglen.studio

Paglen avait fait l'expérience directe de ces biais avec l'artiste Hito Steyerl et la compositrice et musicienne Holly Herndon<sup>627</sup>, avec ses *Machine-Readable* (2017). Concernant Hito Steyerl, selon ses expressions faciales, la détection du genre (déjà sujette en elle-même à discussion) ne renvoie pas les mêmes résultats, suggérant par la même des biais algorithmiques.

Ce développement de la reconnaissance faciale et, par extension, de toutes les technologies biométriques va de plus en loin, notamment en ce qui concerne la détection des émotions. Shoshana Zuboff en situe les origines à deux périodes. La mise au point, en 1978, du Facial Coding System (FAC) par les psychologues Paul Ekman et Wallace Friesen<sup>628</sup>. Il s'agissait d'une méthode de description des mouvements du visage, qui listait six émotions basiques (colère, peur, tristesse, joie, dégoût et surprise) et comportait vingt-sept unités d'action faciale. Puis, en 1997, l'émergence du « concept d'informatique affective » par la publication de l'ouvrage de l'ouvrage *Affective Compu-*

627 Holly Herndon est une compositrice et musicienne expérimentale. Elle travaille beaucoup avec la voix et les outils informatiques. Elle explore actuellement les relations entre l'intelligence artificielle et la musique, l'apprentissage profond (en particulier le *Vocal Deep Fake*). Dans son album *PROTO* (2019), certains chœurs sont générés artificiellement à partir d'autres. Ses derniers clips mettent en scène ses recherches visuelles. Elle a même récemment créé son double (un *Digital Twin*, nommé Holly+) afin d'explorer tous les potentiels de ces nouveaux territoires musicaux. Pour elle, ces technologies doivent être source d'appropriation pour les artistes, pour créer des sonorités inédites ou des manières différentes de composer.

628 Shoshana Zuboff, *op. cit.*, p. 385.

*ting*<sup>629</sup>, qui documentait le travail de la chercheuse Rosalind Picard au MIT. Celle-ci cherchait à restituer les émotions conscientes et inconscientes grâce à l'analyse du visage. La démarche nécessitait de récupérer des données, et Picard mettait déjà à l'époque en garde contre l'utilisation abusive de ces données. Zuboff signale pourtant que bon nombre de ces chercheurs en informatique affective finissent par travailler pour de grosses compagnies<sup>630</sup>, qui sont sans cesse en demande de « cartographies des émotions ». Ainsi la chasse à la voix<sup>631</sup> n'est pas isolée, elle est l'un des maillons d'une traque plus étendue de visages, de gestes, de comportements.

Cette combinaison d'intelligence artificielle a été formée à isoler, puis à capter et à restituer les comportements les plus subtils, les plus intimes, d'un clin d'œil involontaire à une mâchoire qui s'affaisse quelques secondes sous l'effet de la surprise. Combinés à des logiciels, ces capteurs reconnaissent et identifient des visages et mesurent des distances spécifiques qui leur permettent d'interpréter des « micro-expressions », des mouvements d'yeux, des émotions, des humeurs, de la tension, de la dissimulation, de l'ennui, de la confusion, des intentions et bien plus : tout cela en temps réel<sup>632</sup>.

Ce qui se joue de manière encore plus indicible est pourtant en arrière-plan, dans l'indexation et l'interprétation de ces données. Pour Jean-Paul Fourmentraux, ces « empreintes digitales visuelles » permettent à la fois de qualifier, mais aussi de disqualifier. Il s'agit au fond de chercher des comportements suspects, déviants, avec toute la part de stigmatisation que cela peut engendrer<sup>633</sup>.

---

629 Rosalind Picard, *Affective Computing*, Cambridge, MIT Press, 1997.

630 *Ibid.*, p. 387.

631 Voir la section précédente sur les assistants vocaux, section 3.2.1.

632 Shoshana Zuboff, *op. cit.*, p. 382.

633 Jean-Paul Fourmentraux, *antiDATA, la désobéissance numérique. Art et hacktivisme technocritique*, Dijon, Les presses du réel, coll. « Perceptions », 2020, p. 49.

### 3.2.3 La détection sonore et visuelle comme dispositif de composition

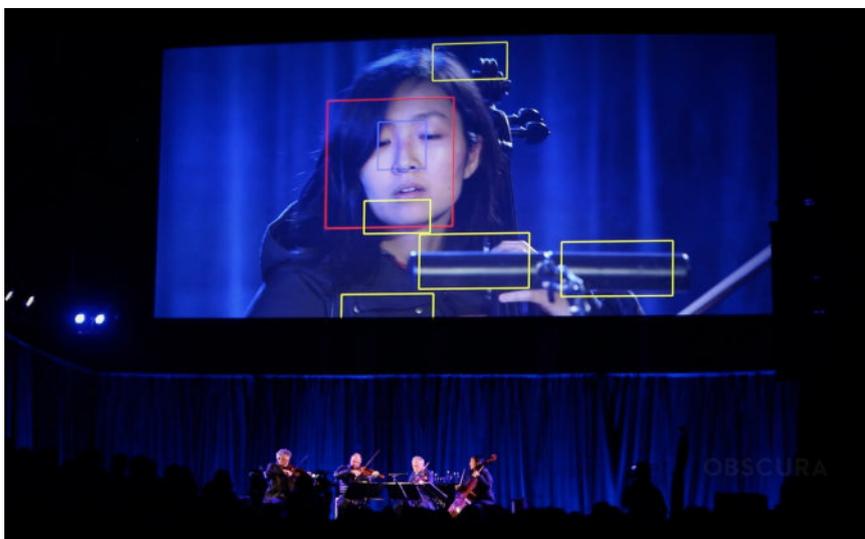


Fig. 158 : Trevor Paglen et le Kronos Quartet, *Sight Machine*, performance audiovisuelle (2018)  
© Cantor Center for Visual Arts at Stanford University

La reconnaissance faciale et les intelligences artificielles sont communément associées à la surveillance. Dans plusieurs de leurs œuvres, les artistes Trevor Paglen et Alexander Schubert créent des mises en abyme de ces processus, dans le contexte de performances scéniques, avec des musiciens.

#### ***Trevor Paglen. Les machines de vision sur scène***

Avec *Sight Machine* (2017)<sup>634</sup>, Trevor Paglen questionne la reconnaissance faciale dans un contexte particulier, celui de la musique et de la scène. Cette œuvre est une performance au cours de laquelle le Kronos Quartet<sup>635</sup> va jouer des pièces de différents compositeurs et compositrices : Laurie Anderson, Steve Reich et Terry Riley. Pendant la performance, une intelligence artificielle va détecter les visages des interprètes, les détourner, et les retransmettre en direct sur un grand écran placé sur la scène. Les visages sont entourés d'informations statistiques concernant leur identité probable et leur humeur.

Le public est donc doublement invité à vivre l'expérience du concert et à suivre les opérations par lesquelles une « machine de vision » (*Sight Machine*) apprê-

<sup>634</sup> Une captation est disponible en ligne sur le YouTube du Holland Festival, <https://www.youtube.com/watch?v=5mg3MXETfj4>, consulté le 30/01/23.

<sup>635</sup> Un quatuor renommé de musique nouvelle et contemporaine.

cie et évalue tout ce qui se passe parmi les musiciens. La vision de ces images projetées peut s'avérer troublante, les machines ne se contentant pas de percevoir. Elles opèrent également des interprétations, traitements et classements par l'intermédiaire de l'IA [...]. Cette utilisation d'algorithmes de détection, associée à des systèmes de surveillance avancés, met au jour la distance ou l'écart entre la perception humaine et celle des machines. Est-il vraiment possible de quantifier l'inspiration, l'expression ou l'émotion humaines ?<sup>636</sup>

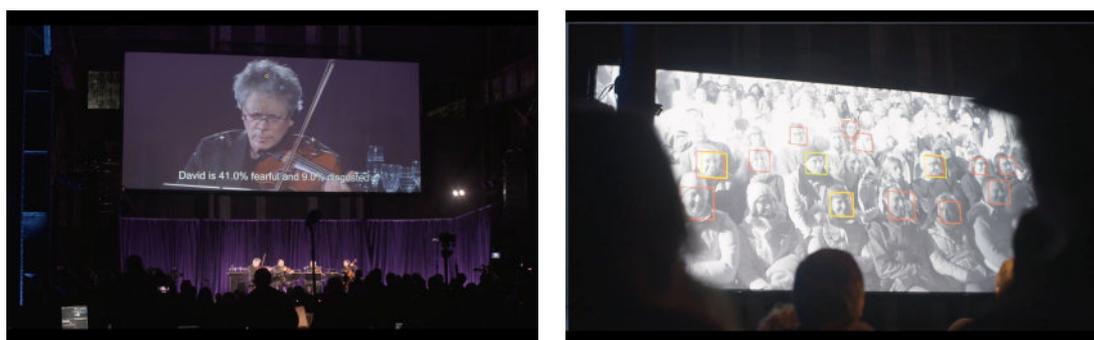


Fig. 159 : Trevor Paglen et le Kronos Quartet, *Sight Machine*, performance audiovisuelle (2018)  
© Barbican

Dans la performance de Paglen, le système accordait à l'une des interprètes un pourcentage « erroné » sur le genre (féminin ou masculin), qui interroge une nouvelle fois la fausse neutralité des algorithmes. Le public lui-même, au cours de la performance, se découvre filmé, puis soumis à cette reconnaissance faciale (cf. image de droite ci-dessus).



Fig. 160 : Trevor Paglen, *Image Operations Op.10* (2018)  
© Trevor Paglen / paglen.studio

636 Jean-Paul Fourmentaux, *antiDATA, la désobéissance numérique, op. cit.*, p. 51.

Paglen emploie le même système dans une pièce vidéo, *Image Operations Op.10* (2018), mais le système va plus loin, en essayant de décrire par des phrases ce qu'il « voit » dans l'image. Par exemple « Une femme tenant un téléphone portable dans sa main ». Certaines descriptions restent encore imprécises (il s'agit en l'occurrence ici d'une femme jouant du violon en gros plan), mais l'utilisation des mêmes dispositifs de reconnaissance que pour les voitures autonomes ou les missiles guidés dans ce contexte crée un décalage fort.

### **Alexander Schubert. Fusions détectionnelles**



Fig. 161 : Alexander Schubert, *Convergence* (2020-21)  
Version scénique  
© Alexander Schubert / alexanderschubert.net

Le procédé employé par Trevor Paglen est poussé encore plus loin — cette fois-ci à de fins visio-musicales — par le compositeur contemporain Alexander Schubert dans sa pièce *Convergence* (2020-2021). Ce dernier y utilise les nouvelles technologies de détection de l'image et du son, mais aussi des intelligences artificielles qui lui permettent d'explorer des manières singulières et novatrices de composer.

La pièce est une performance scénique<sup>637</sup>, entre musique et théâtre, interprétée par un ensemble de musique contemporaine (l'Ensemble Resonanz, constitué de quatre musiciens). La narration audiovisuelle va consister en une mise en abyme, celle de l'utilisation d'une intelligence artificielle pour la composition. Dans l'espace obscur de la scène, l'ensemble n'est pas tout de suite au complet. Les interprètes arrivent un par un sur la scène, et font, chacun leur tour, l'objet d'une analyse faciale, gestuelle et vocale. Lorsqu'un interprète se place au milieu de la scène, une voix off, artificielle, lui donne des ordres. À la manière d'un criminel qui se voit photographié par la police de face, puis de profil, une série de sollicitations débute : différentes postures (assis, debout...), différents jeux instrumentaux et différentes émotions sont demandés par la machine. L'écran nous montre toutes les analyses sous forme de grille, de paramètres, de visages passés au crible. Une fois réuni sur scène, l'ensemble commence à jouer. L'écran donne à voir le processus de détection et de génération qui se met en place.

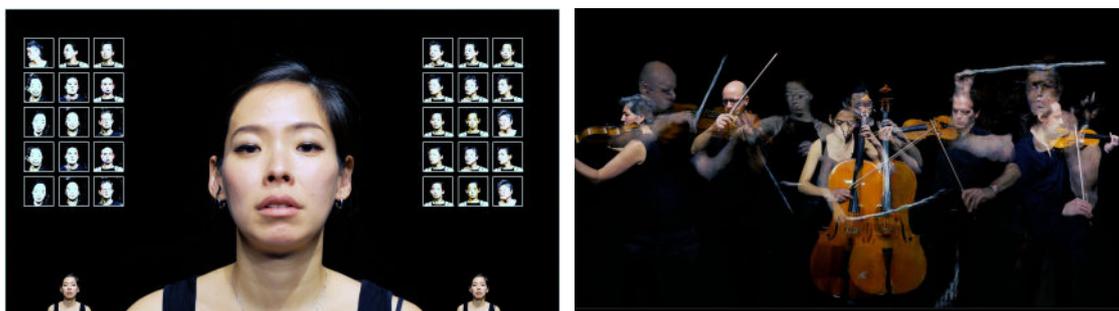


Fig. 162 : Alexander Schubert, *Convergence* (2020-21)  
Version streaming  
© Alexander Schubert / alexanderschubert.net

Tour à tour, ou en même temps, les gestes des instrumentistes sont recalculés, réinterprétés par la machine pour devenir sur l'écran de projection des images fantomatiques, mi-humaines, mi-artificielles des corps. Côté musical, un son de violon devient un cri étrange, une voix chantée devient un son de corde. Les va-et-vient entre scène et écran semblent s'entremêler. Des morphings entre visages et corps se mettent en

637 Deux versions de la pièce existent. L'originale est la version scénique, présentée pour la première fois en Allemagne, à Bonn, en octobre 2020. La seconde est une version en streaming, diffusée à partir de 2021. Dans la version scénique, l'ensemble joue en direct, tandis qu'est diffusée sur un écran en arrière-plan une vidéo. La version streaming est un montage vidéo linéaire que nous détaillerons plus loin.

place, générant un ensemble-fantôme qui accompagne l'original. Les sons humains et les sons d'instruments fusionnent, les images de personnes et les instruments aussi. Il y a un parallèle assez net entre cette pièce et la performance *Violin Power* (1978)<sup>638</sup> de l'artiste Steina, mais elle s'inscrit également dans un sillage plus contemporain, par exemple celui de la performance *Ultrachunk* (2018)<sup>639</sup> de Memo Atken et Jennifer Walshe.

La version streaming<sup>640</sup> est tout aussi troublante, puisqu'elle mêle des captures du réel et des transformations artificielles, mais dans le contexte d'une vidéo linéaire. Plutôt qu'une captation de la version scénique, qui aurait été trop peu lisible (de la vidéo dans la vidéo), Schubert a remonté les séquences, en alternant des captures et des extraits de la vidéo. Dans les deux cas, ce que Schubert donne à voir et à entendre résume l'exploration des procédés utilisés : il s'agit d'une composition en dialogue avec son propre processus de fabrication. Les technologies utilisées par Schubert ont un objectif : arriver à générer de nouvelles entités vidéos et de nouveaux sons, à partir de sons et d'images capturées du réel.

Une partie du travail se fait en amont, dans le paramétrage des intelligences artificielles. Pour cela, Schubert utilise une même nature de technique pour l'audio et pour la vidéo : le *Video Deep Learning* et l'*Audio Deep Learning* (à l'aide d'un logiciel développé pour l'occasion avec une équipe de l'Ircam et une autre équipe, de Heldenkom-

638 Nous avons parlé de cette performance dans la deuxième partie. Dans celle-ci, la captation filmée de Steina en train de jouer du violon était modifiée en direct par une manipulation vidéo opérée par le son capturé. Ici Schubert détecte, capture et injecte les images et les sons, créant une performance doublée par ses propres fantômes dynamiques.

639 En 2018 l'artiste Memo Atken et la performeuse Jennifer Walshe ont travaillé sur une performance nommée *Ultrachunk* dont l'une des composantes est une intelligence artificielle, GRANNMA (Granular Neural Music and Audio), créée par Atken et basée sur l'apprentissage automatique. Chaque jour, pendant une année, la performeuse a improvisé vocalement devant sa webcam. Les enregistrements quotidiens de sa voix et de son visage ont servi de matériau à Memo Atken pour entraîner une intelligence artificielle. Une fois que celle-ci est suffisamment nourrie d'informations vocales et iconiques, elle devient capable de générer en temps réel un flux audiovisuel qui recrée une performance sur la base du matériel enregistré. Pendant la performance, Jennifer Walshe improvise vocalement, en duo avec son double artificiel contrôlé par Atken. La représentation spectrale du son est présente sur les deux côtés de la vidéo générée. Il existe également une version installation, autonome et générative. La machine, à partir du matériau, va générer des images et des sons qui n'existaient pas, mais qui pourraient potentiellement être. Comme pour Schubert, le résultat est très étrange, sorte de flux-morphing saccadé qui semble pouvoir s'exercer à l'infini, sans début ni fin, dans laquelle le visage et la voix de la performeuse semblent perdus et enfermés (Une documentation vidéo est disponible en ligne sur le Vimeo de Memo Atken, <https://vimeo.com/314016795>, et une captation performance sur son site, <https://www.memo.tv/works/ultrachunk/>, consultés le 30/01/23).

640 La vidéo streaming est disponible en ligne sur le site d'Alexander Schubert, <https://www.alexanderschubert.net/works/Convergence.php>, consulté le 30/01/23.

binat Technologies, pour la vidéo). Le scientifique Jean-Claude Heudin définit le *Deep Learning* (ou « apprentissage profond ») comme « une famille d'algorithmes d'apprentissage (*Machine Learning*) pour entraîner des réseaux de neurones composés de plusieurs couches internes et potentiellement un grand nombre<sup>641</sup> ». Il explique que l'on utilise le terme « profond » pour évoquer la « profondeur » du réseau, c'est-à-dire son nombre de couches, et que si les chercheurs s'y intéressent depuis 2006, ce n'est que dans les années 2010 que les choses se sont développées<sup>642</sup>. À un moment du « scanage », la voix prononce d'ailleurs ces mots : « The more we scan, *deeper* is the result ».

Les deux systèmes, audio et vidéo, sont séparés, mais ils utilisent les mêmes principes. Il semble que toute la partie vidéo soit entièrement préparée en amont, car elle nécessite plus de ressources informatiques, mais que la partie audio puisse être effectuée en temps réel, à la condition qu'elle ait été « nourrie » suffisamment.



Fig. 163 : Capture de l'interface du logiciel d'apprentissage audio, développé avec l'Ircam pour *Convergence*  
© Alexander Schubert / alexanderschubert.net

Pour pouvoir travailler, le *dataset* (la base de données qui nourrira le système de génération sonore) doit être structuré. Un protocole est donc établi — même si des sessions d'improvisation sont aussi prévues. Dans le cas de l'audio et du logiciel développé par

<sup>641</sup> Jean-Claude Heudin, *Comprendre le Deep Learning : une introduction aux réseaux de neurones*, Paris, Science eBook, version Kindle, 2016, p. 130.

<sup>642</sup> *Ibid.*, p. 120.

l'Ircam, le système comporte trois parties principales : l'*Input* (l'entrée), le *Latent Space* (l'espace latent), l'*Output* (la sortie).

En amont, Schubert fait enregistrer à ses interprètes différentes natures de motifs : un motif musical de corde réalisé de manière conventionnelle, un autre motif réalisé de manière plus expérimentale (frottements, percussions...), un chant, un parlé, un cri. Ces captures sont appelées les entrées (INPUT). Elles sont numérisées, non pas de manière traditionnelle, mais de manière paramétrique. La machine, entraînée à être nourrie d'une certaine nature de données, compose un modèle abstrait, par exemple celui d'un son de violon. La base de données contient alors différents modèles paramétriques de différentes natures de sons.

Par exemple, si un son de voix (par exemple la prononciation d'un prénom) est entré (*Input*), il est traduit paramétriquement dans le *Latent Space* (par une sorte de réduction à des ensembles de nombres), et peut être reconstruit en sortie (*Output*) en utilisant un autre modèle figurant dans la base ; celui du violon par exemple. Auquel cas, les fluctuations de la voix seront reconstruites en fluctuations de violon (réalisé de manière conventionnelle) généré artificiellement.

Il s'agit de techniques fondées sur les Auto-encodeurs et les GAN (*Generative Adversarial Network* ou Réseau antagoniste génératif). Ceux-ci permettent un encodage et un décodage par transfert. Les systèmes à auto-encodeurs sont définis ainsi par Jean-Claude Heudin :

Il s'agit d'un type de réseau particulier comportant une couche d'entrée et de sortie comprenant un même nombre de neurones, mais avec une couche cachée plus restreinte. De cette manière, le réseau « encode » ses entrées en réduisant l'information. En entraînant un autoencodeur sur un volume important de données et où les sorties désirées sont les données d'entrée, on obtient de bons « extracteurs » de caractéristiques propres au jeu de données présent<sup>643</sup>.

On peut donc créer des interchangeabilités, reconstruire une séquence de jeu de violon en séquence de cri, une séquence de chant en séquence de percussion, et ainsi de suite. Il est également possible, dans le *Latent Space*, d'appliquer des transformations, des interpolations, du morphing.

---

643 Jean-Claude Heudin, *Comprendre le Deep Learning, op. cit.*, p. 128-129.

La nature de la technologie utilisée ici est la même pour l'audio et la vidéo. Il y a un traitement par dérive très semblable, mais les processus ont été pris en charge par deux équipes différentes. Les apprentissages audio d'une part, ceux vidéo d'autre part. Ils sont ensuite rassemblés par un montage. Nous avons demandé à Philippe Esling de l'Ircam, qui a travaillé sur la partie audio, de nous dire si des dérives (transferts) auraient pu être réalisées entre de l'audio en *Input* et de la vidéo en *Output* (ou inversement) ; cela lui semble tout à fait possible, mais n'a pas été mis en œuvre ici. Probablement le trouble est plus grand (et la compréhension meilleure pour le spectateur) quand l'image générée reste liée et confrontée directement à son original (idem pour le sonore). Le processus d'une image reconstruite en son restera plus opaque à comprendre.

On gravite ici, avec ces techniques, entre la capture du réel et la synthèse, dans des univers fantômes. On peut parler de transformation, de transfert, de dérive. Le son artificiel, ou l'image artificielle reste liée à son original, dans une « inquiétante étrangeté » (le terme anglais *uncanny* est souvent utilisé pour définir une telle nature d'image).

À travers cette expérience, Schubert tente de comprendre, d'explorer les manières dont la machine voit et entend. Vers la fin de la performance, la voix prononce ces mots : « I'm also just a mirror for you. As I provide a mirror for you<sup>644</sup> ».

### 3.2.4 Du visage à la voix, de la voix au visage : clones et chimères numériques

Jacques Aumont nous le rappelle : dans le contexte du cinéma, la voix et le visage sont tous deux des vecteurs fondamentaux de l'identité d'une personne à l'écran<sup>645</sup> ». Il pointe à ce propos les paradoxes du doublage et les biais déjà en jeu à l'époque du parlant. Désormais, comme nous l'avons vu dans ce chapitre, la voix et le visage sont soumis aux mêmes processus d'apprentissage et de reconnaissance automatique. En termes purement biométriques, la synchronisation visage-voix emblématique du cinéma parlant a muté. Elle est en passe de devenir le lieu d'intersection d'un « faisceau

644 Alexander Schubert — *Convergence* [Ensemble Resonanz] @Kampnagel/Eclat, <https://www.youtube.com/watch?v=o5UXkIWJciQ>, consulté le 30/01/23.

645 « [...] deux voies royales qui me permettent d'accéder à l'humanité de l'autre homme » (Jacques Aumont, *Le visage au cinéma*, op. cit., p. 52).

d'indices » permettant d'établir l'identité et les émotions d'un individu. Que ce soit à l'échelle de dispositifs militaires, gouvernementaux ou liés aux grandes entreprises du numérique, les deux modalités, faciale et vocale, seront, il nous semble, de plus en plus croisées. Shoshana Zuboff le résume ainsi : « Il fut un temps où vous exploriez grâce à Google. Maintenant, c'est vous que Google explore<sup>646</sup> ».

Ainsi, la biométrie<sup>647</sup> vocale (reconnaître l'identité d'un timbre de voix) est en plein essor, dans l'idée de la croiser avec la biométrie faciale pour obtenir des résultats encore plus pertinents. On peut ainsi lire, sur les sites web d'entreprises de consulting, ce type de publicité qui vante les bienfaits de systèmes technologiques de reconnaissance vocale :

Les algorithmes sont désormais capables d'identifier plus de 1000 caractéristiques de la voix (spectre, fréquence, timbre, débit, accent, silence, ton, rythme...). Cette technologie permet d'identifier un individu même lorsque sa voix subit des variations liées aux conditions d'enregistrement, à l'émotion, à la santé (exemple : un rhume)<sup>648</sup>.

La détection et l'analyse ne sont plus les seules manières d'appréhender les relations entre voix, visage et numérique. Couplés aux technologies d'apprentissage des intelligences artificielles, il est devenu possible de générer de façon très réaliste des visages, mais aussi des voix, à partir de modèles ou de manière totalement artificielles, en fusionnant des visages à partir d'une base. Tous ces procédés croisés vont donner lieu à de nouvelles formes d'interrelation entre le visage et la voix, dont les territoires s'étendront au probable, à l'anticipation et au virtuel. Marion Zilio l'analyse déjà à propos de potentialités de recherche d'un visage probable : « Le datamining, la traçabilité et le profilage sont devenus des outils d'anticipation, où le probable recouvre désormais le potentiel<sup>649</sup> ».

---

646 Shoshana Zuboff, *op. cit.* p. 354.

647 L'ensemble de procédés technologiques visant à la reconnaissance d'un individu.

648 « L'usage de la biométrie vocale au sein du secteur bancaire », Thinkmarket, <https://www.thinkmarket.fr/uploads/2020/09/oeil-sur-biometrie-vocale-VF3.pdf>, consulté le 30/01/23.

649 Marion Zilio, *op. cit.*, p. 147.



Fig. 164 : Heather Dewey-Hagborg et Chelsea Manning, *Probably Chelsea* (2017)  
© Benoît Montigné, Biennale Némò, octobre 2021

Une œuvre de l'artiste Heather Dewey-Hagborg, *Probably Chelsea* (2017) reflète bien cette idée. Pour celle-ci, Chelsea Manning, la lanceuse d'alerte, a confié à l'artiste des échantillons de salive et de peau afin que cette dernière, grâce à des systèmes utilisés de manière expérimentale par la police scientifique, génère les visages potentiels. Trente visages « probables » ont ainsi été générés et modelés. Ils sont suspendus dans l'espace installé.

L'autre technique contemporaine est celle du *Deep Fake*, qui permet, à partir de « réseaux antagonistes génératifs » (en anglais les GAN, Generative Adversarial Network) de générer des visages artificiels, voire des vidéos avec le visage et la voix en rendu réaliste. En 2019, l'artiste Bill Posters avait ainsi généré de fausses vidéos de plusieurs célébrités, de l'artiste Marina Abramović à Mark Zuckerberg, le fondateur de Facebook. Le rendu était si réaliste que les vidéos sont devenues virales. Le système technologique entremêle un rendu de voix et un rendu de visage synchronisé, avec toujours des cadrages en gros plan ou en plan rapproché poitrine<sup>650</sup>. De véritables portraits parlants, mais aussi des marionnettes à qui l'on peut faire dire ce que l'on veut.

<sup>650</sup> Voir au premier chapitre de cette partie, section 3.1.1.



Fig. 165 : Bill Posters, *Big Dada/Public Faces* (2019)  
Séquence avec Marina Abramović  
© billposters.ch

Il devient également envisageable d'inférer un visage à partir d'une voix<sup>651</sup>, c'est-à-dire de générer un visage type. Des intelligences artificielles peuvent opérer des lectures labiales sur des vidéos, afin de détecter un échange vocal, reconstituer une voix à partir d'un visage, générer un modèle de voix à partir d'un échantillon de quelques minutes d'une voix, animer automatiquement un visage à partir d'une voix réelle ou artificielle.

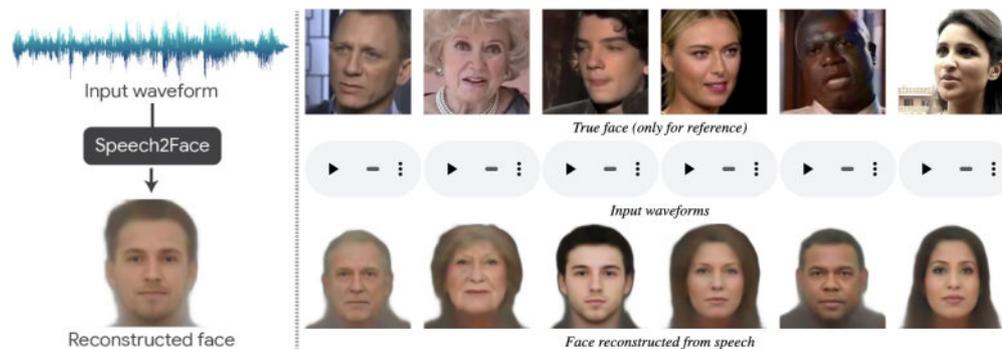


Fig. 166 : Exemple de technologie Speech2face  
Les portraits du bas sont générés à partir d'échantillons de voix.  
Ceux du haut sont des captures des modèles originaux.  
© speech2face.github.io

<sup>651</sup> *Speech2face*, un exemple de recherche : Tae-Hyun Oh, Tali Dekel, Changil Kim, Inbar Mosseri, William T. Freeman, Michael Rubinstein, Wojciech Matusik, « Speech2Face: Learning the Face Behind a Voice », 2019, <https://speech2face.github.io/>, consulté le 30/01/23.

Autre exemple parmi d'autres, l'annonce suivante, publiée par un institut de recherche en sciences informatiques du CNRS, propose un postdoctorat en « inférence du visage à partir de la voix » :

Il ne s'agira pas de générer le visage exact correspondant à la voix, mais de générer un visage qui permette de ressortir les caractéristiques discriminantes principales du visage (genre, âge, ethnicité, attributs cranio-faciaux). Ceci peut présenter de nombreux intérêts dans le cadre d'applications liées à la sécurité. En effet, inférer un visage à partir d'une voix peut permettre ensuite à un opérateur d'effectuer différentes tâches à partir de l'image de visage inférée. Nous pouvons citer par exemple : la vérification de la cohérence entre une voix et un visage, la recherche du visage inféré dans une base de données<sup>652</sup>.

Dans le registre des émotions, des chercheurs de l'Ircam travaillent actuellement sur les « transformations vocales émotionnelles<sup>653</sup> », c'est-à-dire l'idée de travailler sur les indices d'expression vocale comme la hauteur ou le timbre à l'aide de filtres, similaires à ceux utilisés pour les visages. Parmi les exemples d'application, « celui de l'employé de call-center (centre d'appels) qui transforme la voix de ses clients mécontents pour moins souffrir de leur agressivité ; du politicien qui rend sa voix plus convaincante ; mais aussi des situations thérapeutiques comme celle d'un patient dépressif qui rend sa voix plus "souriante" pour ses proches, ou d'une personne stressée qui se calme en écoutant une version apaisée de sa propre voix<sup>654</sup> ».

D'autres exemples pourraient être cités, de la lecture labiale<sup>655</sup> automatisée à la diarisation des orateurs<sup>656</sup>. Quelle que soit la nature des intentions derrière ces technologies à double tranchant, les nouvelles modalités de traitement et d'analyse de la voix

---

652 Offre de Post-doctorat « Inférence du visage à partir de la voix », GdR ISIS (Information, Signal, Image et ViSion), structure d'animation de l'INS2I (Institut des Sciences Informatiques et de leurs Interactions) du CNRS, <https://www.gdr-isis.fr/index.php/annonce/8695/>, consulté le 11/08/2022.

653 Nadia Guerouaou, « Rendre sa voix plus souriante : deepfakes et filtres vocaux émotionnels », AOC, juillet 2022, <https://aoc.media/analyse/2022/07/05/rendre-sa-voix-plus-souriante-deepfakes-et-filtres-vocaux-emotionnels/>, consulté le 30/01/23.

654 Jean-Julien Aucouturier, Nadia Guerouaou et Guillaume Vaiva, « L'étonnante acceptabilité des deepfake », Le journal CNRS, janvier 2022, <https://lejournel.cnrs.fr/billets/letonnante-acceptabilite-des-deep-fake>, consulté le 30/01/23.

655 La lecture sur les lèvres, utilisée par les personnes sourdes et malentendantes.

656 La diarisation consiste à arriver à séparer les locuteurs dans un flux audio unique mêlant des conversations.

et du visage par le numérique semblent dans tous les cas converger, et constituer un terrain fertile d'exploration artistique.

### **3.3 Choralités machiniques**

#### **3.3.1 Chœurs virtuels. Opéra d'ordinateurs**



Fig. 167 : Extrait documentation vidéo Nahas & Huitric (1993)  
© Archives vidéo de Paris 8

Loin des utilisations diagrammatiques du visage et de la voix décrites dans le chapitre précédent, les recherches sur les modélisations du visage et de la voix ont aussi pu emprunter des chemins plus poétiques. C'est le cas du travail pionnier de Monique Nahas et d'Hervé Huitric autour des chœurs virtuels. Débuté en 1990, il s'est développé pendant plusieurs années au sein du département Arts et Technologie de l'Image (ATI) de l'Université Paris 8.

Après quelques premières expérimentations en 1988 sur la modélisation du visage en 3D, Monique Nahas et d'Hervé Huitric ont progressivement amélioré le réalisme de leurs créations. Mais dès le départ, ce réalisme ne fut pas au service d'une représentation fidèle, mais plutôt fut plutôt le fruit d'expérimentations utilisant le visage-voix comme pinceau, pouvant démultiplier ses occurrences, ou bien les fusionner.



Fig. 168 : Monique Nahas et Hervé Huitric, *Masques et bergamasques* (1990)  
© Archives vidéo de Paris 8

Dans *Masques et bergamasques* (1990)<sup>657</sup>, les deux artistes-chercheurs mettent en scène des portraits parlants en trois dimensions. Dans l'une des séquences, un visage (numérisé en 3D, dont le modèle semble être Monique Nahas), tout en parlant, commence à se démultiplier, à la manière d'une mitose cellulaire. La voix, elle aussi, suit le même processus. Quelques dizaines de secondes plus tard, ce sont trois entités qui parlent, dans ce qui est devenu un chœur vocal virtuel. Ce principe sera au cœur des différentes expérimentations et installations dont nous allons ici parler.

Progressivement, avec la synthèse vocale, les entités modélisées acquièrent plus d'autonomie. À partir de l'année 2000, Nahas et Huitric expérimentent ce qu'ils nomment des « conversations automatiques » (elles peuvent aussi être des monologues). Les protagonistes virtuels (les « causeurs ») peuvent communiquer automatiquement entre eux, grâce au programme Elise, inspiré du bot Eliza. Une synchronisation en réseau est possible sur autant d'ordinateurs que souhaité. Les causeurs peuvent échanger des commentaires, des chants, des poésies. Ils peuvent parler ou chanter ensemble. Une session d'échanges peut durer jusqu'à quarante-huit heures.

<sup>657</sup> Une version est disponible en ligne sur le site des archives vidéo de Paris 8, <https://www.archives-video.univ-paris8.fr/video.php?recordID=305>, consulté le 30/01/23.



Fig. 169 : Atelier de chœur virtuel. « Les visages chantent les voyelles d'un texte lu par un programme de synthèse vocale ». Un bâton, le flock of bird, permet, grâce à ses mouvements de démarrer les chants, de répartir les visages. Des commandes sont aussi disponibles sur l'ordinateur.

© Monique Nahas et Hervé Huitric / huitric-nahas.org

En 2002, le concept de chœur virtuel est déposé. Il se compose d'un réseau d'ordinateurs au sein desquels des visages chantent et parlent. L'ensemble est interactif et contrôlable grâce à un bâton et des commandes sur ordinateur.

Celui-ci [le chœur virtuel] est constitué par un ensemble d'ordinateurs pouvant parler ou chanter en temps réel. Ils doivent être en réseau et leur nombre n'est pas limité. Chaque écran constitue un ou plusieurs chanteurs virtuels. Un ordinateur particulier assiste le « chef d'orchestre » : avec ses capteurs, il lui offre une baguette aux dimensions multiples. À titre d'exemple, nous avons utilisé un « flock of bird » pour déclencher des sons, des transformations de sons et d'images, et nous nous sommes servis de joystick, de menus, par exemple pour des déplacements d'un écran à l'autre, et diverses déformations. Bien évidemment on peut faire une combinatoire de tout ce qui peut se trouver comme capteurs, la seule limite restant toujours de ne pas créer une complexité trop difficilement gérable en temps réel<sup>658</sup>.

658 Monique Nahas et Hervé Huitric, « Chœur virtuel », <http://www.huitric-nahas.org/installations/installations2.htm>, consulté le 30/01/23.



Fig. 170 : *Comme cela vous chante* (2002)

Villette Numérique

© Monique Nahas et Hervé Huitric / huitric-nahas.org

Le chœur devient peu à peu un opéra expérimental interactif avec *Comme cela vous chante* (2002), qui utilise douze ordinateurs, ou performance multimédia avec *Les Glossolalies d'Orphée* (2002), avec cinq ordinateurs. Une description de *Comme cela vous chante* est donnée cette année-là dans le Monde Interactif.

Sur les écrans d'ordinateur, des visages en images de synthèse se déplacent et chantent dans un univers organique en trois dimensions. Un chef d'orchestre les dirige à l'aide d'une étrange baguette. Munie d'un capteur magnétique, celle-ci permet de modifier les expressions des différents visages, d'en moduler les voix et de déclencher ou geler les animations de fond en 3 D. Mi-électronique, mi-humain, leur chant, initié à partir d'une phrase anodine prononcée dans un microphone par un spectateur, pétrifie. Cette phrase est d'abord répétée par un ou plusieurs « visages », puis un programme spécifique l'analyse en temps réel et en extrait une suite de voyelles sonores qui s'épanouissent en un véritable chant. La voix de soprano d'une chanteuse bien réelle s'élève pour accompagner la complainte mystique de ce chœur composé de machines<sup>659</sup>.

Plusieurs techniques singulières sont utilisées, citons celle du pinceau 3D, qui consiste à utiliser un visage parlant comme « brosse », comme on le ferait dans un logiciel de dessin. En appuyant et déplaçant un curseur à l'écran, on clone en temps réel le vi-

659 Extrait d'un article du Monde Interactif de mars 2002, <http://www.huitric-nahas.org/installations/installations4.htm>, consulté le 30/01/23.

sage-voix sélectionné. Le son de la parole est lui aussi cloné, ce qui crée, selon les tempos de décalage et de déphasage, différentes natures d'échos<sup>660</sup>.



Fig. 171 : Extrait documentation vidéo Nahas & Huitric (1993)  
© Archives vidéo de Paris 8

Monique Nahas et Hervé Huitric, tout en développant et en utilisant des technologies au rendu réaliste, les ont détourné pour créer des opéras interactifs. Eux-mêmes le résumant ainsi.

Une approche commencée depuis des années et qui se poursuit, autour des visages virtuels parlants et chantants, capables de conversations et d'expressions, d'émotions avec des êtres réels aussi bien que virtuels. Nous avons réalisé plusieurs installations interactives : conversation entre un intervenant réel et un réseau de visages virtuels, conversation entre deux acteurs virtuels se répondant, duo entre un chanteur réel et un chanteur virtuel, canon d'un chœur virtuel déclenché par la parole d'un spectateur. Les dimensions de l'action, de la perception, de la sensation et de l'émotion sont partie intégrante de nos recherches comme en témoignent aussi les travaux réalisés par Monique Nahas sur la reconnaissance des expressions chimériques, ainsi que les travaux bien antérieurs sur l'expression des émotions d'un visage de synthèse<sup>661</sup>.

660 Des documentations vidéo sont disponibles en ligne sur le site des archives vidéo de Paris 8, l'une autour de l'animation des visages à l'aide de capteurs, <https://www.archives-vidéo.univ-paris8.fr/video.php?recordID=315>, l'autre de la digitalisation par technique Laser, <https://www.archives-vidéo.univ-paris8.fr/video.php?recordID=314>, consultés le 30/01/23.

661 Hervé Huitric et Monique Nahas, *Films d'images de synthèse (1983-1994)*, Bande-annonce, <https://octaviana.fr/document/FVAONUMBA11#?c=&m=&s=&cv=>, consulté le 30/01/23.

### 3.3.2 *Modell 5*, de Granular Synthesis. Grains de visage-voix



Fig. 172 : Granular Synthesis, *Modell 5* (1994-1996)  
© Ars Electronica

Le principe de choralité développé par le couple Nahas-Huitric est également au cœur d'une installation vidéo-sonore, créée par le duo Granular Synthesis<sup>662</sup> (Kurt Hentschläger et Ulf Langheinrich). Dans leur œuvre *Modell 5*<sup>663</sup> (1994-1996), il ne s'agit toutefois plus de génération synthétique, mais de la réplique d'une même séquence vidéo, dans une démarche proche du Djing et du Vjing. *Modell 5* a été présentée pour la première fois en 1994 à l'ICA London (The Institute of Contemporary Arts in London). Il existe plusieurs versions de cette pièce, toutes nommées selon une syntaxe issue du *versioning* informatique (*Modell 5.1* jusque *Modell 5.7*)<sup>664</sup>.

#### ***Entre accès programmé et synthèse granulaire***

Que ce soit dans un mode installé ou performé, *Modell 5* se présente sous la forme d'un « tétrapyque » formé de projections verticales juxtaposées<sup>665</sup>, chacune nous donnant à voir un même gros plan en vidéo du visage d'une jeune femme — des clones vidéo en

662 Souvent orthographié *Granular Synthesis*, le nom original du collectif était GRANULAR-SYNTHESIS, en majuscules, avec le signe ~.

663 Une captation vidéo est disponible en ligne sur le Vimeo d'Anti-Utopias, <https://vimeo.com/43744967>, consulté le 30/01/23.

664 Par ailleurs, le titre original variait entre *Model 5*, ou *Motion Control Modell 5*.

665 Chaque projection (habituellement opérée par l'arrière ou directement sur un mur blanc) est à grande échelle (au minimum 3 x 4 m selon la documentation technique) mais doit occuper le maximum de surface.

quelque sorte — dont les expressions faciales vont évoluer au cours de la performance<sup>666</sup>. Les images proviennent d'une captation réelle, filmée, de la performeuse japonaise Akemi Takeya. On se rend très rapidement compte que l'image et le son de la voix sont manipulés par les artistes, de manière très saccadée, parfois violente. Les deux artistes ne sont pas sur scène (il n'y en a pas toujours), mais ils se trouvent dans les rangs du public, ou bien sur le côté, selon les espaces dessinés pour la performance. Leur situation reflète celle d'une régie FOH (*Front Of House* en anglais), c'est-à-dire le groupe de techniciens qui ajustent et paramètrent l'environnement scénique (les lumières, le son...) pendant le spectacle. Le public est donc face aux 4 écrans de grande taille, qui doivent couvrir le maximum d'espace.

Il est important que les écrans forment un mur dans l'espace donné, plutôt que de se trouver dans une situation de type cinéma (un écran sur un mur)<sup>667</sup>.

Nous sommes presque en face d'une installation dont l'image et le son seraient manipulées en live<sup>668</sup>. Ce qui confère également à cette installation-performance une certaine contemporanéité, c'est l'utilisation du format vertical pour le gros plan, assez inhabituel dans les installations écraniques à cette période, si ce n'est chez Bill Viola par exemple. Il est d'ailleurs mentionné plusieurs fois dans la documentation technique « portrait (not landscape) », car les projecteurs doivent être tournés à 90° pour obtenir l'effet.

Côté son, des enceintes semblent être placées derrière ou juste au-dessous des projections, afin de créer une spatialisation ; cela respecte le principe du cinéma muet des premiers temps, où le son était parfois situé derrière l'écran<sup>669</sup>. Chaque portrait-écran

666 Préexistent également d'anciennes versions. Dans ces dernières, un degré de plus dans l'idée de performance live était prévu. Le performeur (qui était alors un homme) filmé pour la vidéo de base était aussi présent dans une pièce connexe à la performance, et était filmé en direct avec un cadre caméra similaire aux versions préexistantes. Ce direct était retransmis sur l'un des quatre écrans. Se jouait donc en quelque sorte un live dans le live, un flux direct manipulé en direct parmi des manipulations de flux préenregistrés.

667 « It's important that the screens become a wall in the given space, rather than a cinema situation (like a screen on a wall) »

668 Il existe d'ailleurs des versions seulement installées, ainsi qu'une version DVD qui adopte un recadrage de l'ensemble, en se focalisant sur le premier écran.

669 Dans le cinéma muet des premiers temps, l'espace situé derrière l'écran pouvait servir à accueillir tout ce qui était de l'ordre de la sonorisation en direct, notamment des bruiteurs ou bien un bonimenteur (un commentateur). Le son semblant sortir de l'écran, cela favorisait pour le spectateur l'ancrage entre l'image et le son (Alain Boillat, *Du bonimenteur à la voix-over, Voix-attraction et voix-narration au*

possède son ainsi propre son. S'ajoute à cela un son complémentaire, celui que les artistes génèrent avec un système d'infrabasse. Cela confère un côté physique à l'expérience déjà émotionnelle.



Fig. 173 : Granular Synthesis, *Modell 5* (1994-1996)  
GS-Retrospective, STRP Eindhoven, 2009  
© Granular-Synthesis

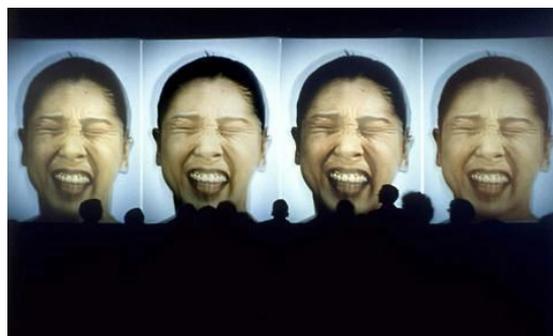


Fig. 174 : Granular Synthesis, *Modell 5* (1994-1996).  
MAK\_Museum of Applied Arts, Vienna, 1995  
© Bruno Klomfar /Gebhard Sengmueller

Les quatre projections sont parfois synchronisées, ou prennent parfois leur indépendance et se comportent différemment afin de créer des contrastes. Les expressions préexistantes de la performeuse — manifestations d'une souffrance dont on ne connaît pas l'origine —, couplées aux manipulations numériques de va-et-vient saccadés, placent le spectateur-auditeur dans une position très inconfortable. Les visages semblent prisonniers, manipulés, secoués, avec quelques moments de calme. Pour décrire ce travail, les artistes emploient le terme *Cyborg*, et effectivement, nous sommes à la frontière du réel et de l'artificiel.

La performance est ancienne, mais elle produit toujours la même émotion. Elle est d'autant plus d'actualité à la lumière des manipulations numériques contemporaines tel que le *Deep Fake*.

### ***L'image-son solidaire***

Dans la plupart des textes consacrés à cette œuvre, la technique qui sert de base à cette installation-performance est décrite comme de la « synthèse granulaire », en anglais *Granular Synthesis*, qui est justement le nom du collectif.

---

*cinéma*, Lausanne, Antipodes, coll. « Médias et histoire », 2007).

La synthèse granulaire, de nos jours répandue, a été théorisée par Dennis Gabor en 1947, mais ne s'est réellement développée qu'avec les ordinateurs.

La synthèse granulaire construit un son à partir de grains ou signaux acoustiques élémentaires de très courte durée, dont la densité est définie par des lois stochastiques. La recherche se poursuit sur la fréquence statistique à donner à chaque « grain »<sup>670</sup>.

En somme, on découpe un signal primaire ou un fragment sonore numérisé préexistant en microéchantillons (les « grains ») qui sont ensuite recombinaés, selon un certain nombre de paramètres (nombre d'échantillons, degré d'aléatoire, modes de recombinaison, etc.). Ce réassemblage donne alors à entendre un résultat totalement différent de la source, un objet qui semble constitué de grains sonores, comme des sortes de pluies ou de vents électroniques.

Le procédé est en réalité utilisé ici de manière moins poussée que dans la véritable synthèse granulaire, mais il donne la même impression de recomposer, en temps réel, la vidéo originale à partir de prélèvements de microfragments.

Il est toutefois important de noter qu'il n'y a pas d'« effets » particuliers mais essentiellement des manipulations temporelles. Dans les faits, cela revient à peu de choses près à emmener la tête de lecture à différents endroits, pendant certaines durées, avec des boucles en aller-retour, comme peuvent le faire des monteurs sur une table de montage avec une molette de défilement. C'est-à-dire un procédé similaire à celui utilisé par l'artiste Steina<sup>671</sup>, une forme de « saut » interactif évolué.

Ce principe de navigation, rendu malléable par le numérique, revêt dans *Modell 5* une importance particulière, car il autorise la manipulation en temps réel à une échelle audiovisuelle fine. Notre perception n'étant plus à même de suivre les coupures, se recrée alors de nouvelles continuités, abstraites, inattendues. Une manipulation extrêmement fine du matériau est alors possible. Cette technique, appliquée aux captations

---

670 Claude Fatus, *Vocabulaire des nouvelles technologies musicales*, Paris, Minerve, coll. « Musique ouverte », 1994, p. 174.

671 Voir le deuxième chapitre de la deuxième partie, section 2.2.3.

vidéos de la performeuse, et manipulée subtilement en temps réel, permettent de à la fois de faire surgir des coupes violentes et des continuités douces.

Dans *Modell 5* s'opère un systématisme, la volonté radicale de manipuler un matériau unique — la captation filmée de la performeuse —, mais en préservant la synchronicité originale de l'image et du son sur chacun des quatre écrans. Pour le public, il y a une tension car, quelles que soient les manipulations temporelles effectuées, de manière lente ou rapide, le son et l'image restent solidaires.

De par l'utilisation de ce procédé, jamais nous n'entendons la voix originale de la performeuse ni les expressions de son visage. Parce qu'ils sont sans cesse saccadés, une artificialité se crée avec ces visages-voix pourtant issus d'une captation réelle.

### 3.3.3 Confessions à des lumières de voix



Fig. 175 : Rafael Lozano-Hemmer, *Voice Array* (2011)  
© Rafael Lozano-Hemmer / lozano-hemmer.com

Comme nous l'avons décrit dans la section précédente, *Modell 5* de Granular Synthesis consiste en un *scratching*<sup>672</sup> audiovisuel qui produit un empêchement permanent : des salves de sauts temporels qui ne permettent jamais à la séquence vidéo, et donc à la voix et aux expressions de visage de se développer naturellement.

---

672 Voir la section 2.2.3 consacrée au saut interactif et au *scratching* dans la deuxième partie.

Dans une démarche en contraste avec celle de *Granular Synthesis*, l'artiste multimédia canadien Rafael Lozano-Hemmer va mettre en scène des choralités désincarnées — sans visage —, et travailler sur une polyphonie de voix plurielles. *Voice Array* (2011)<sup>673</sup> est une installation interactive. Le visiteur pénètre dans un espace obscur, au sein duquel il entend des bribes de voix et observe au mur une série (il y en a au maximum 288) de flashes verticaux de lumière affichés sur des bandes de LED.



Fig. 176 : Rafael Lozano-Hemmer, *Voice Array* (2011)  
© Rafael Lozano-Hemmer / lozano-hemmer.com



Fig. 177 : Rafael Lozano-Hemmer, *Voice Array* (2011)  
© Rafael Lozano-Hemmer / lozano-hemmer.com

Sur la gauche (ou droite selon les espaces) de ce mur se trouve un interphone. Le visiteur peut y enregistrer sa voix, sous la forme de son choix : message, confession, chanson, bruits... Lorsque l'enregistrement est terminé, le son est traduit en lumière, puis mis en boucle. Tous les autres flashes s'interrompent momentanément, puis viennent se repositionner successivement, après ce nouveau flash. On comprend que ce sont les enregistrements précédents qui se décalent pour laisser place au nouveau. Lorsque le plus ancien cherche à se montrer (et surtout à se faire entendre), il n'a plus d'espace ; il sera alors entendu une dernière fois distinctement — les autres voix s'évanouissent le temps de cette lecture —, avec tous les flashes lumineux à l'unisson, puis disparaît. Ce que nous entendons est une trace, la parole anonyme d'une personne qui a été là dans les heures ou jours précédents. La polyphonie de voix, qui a fait mourir un échantillon vocal pour en accueillir un autre, reprend ensuite sa lecture.

<sup>673</sup> Une captation vidéo est disponible en ligne sur le Vimeo de bitforms gallery, <https://vimeo.com/34565650>, consulté le 30/01/23.

Le logiciel qui sert de base à l'interaction est une application dédiée, développée spécifiquement pour l'installation. Son fonctionnement met en relief deux caractéristiques fondamentales de la base de données. D'un côté elle peut accueillir des informations codées et aller les chercher à des moments donnés. De l'autre, elle est limitée par la taille de l'espace de stockage ; si celle-ci est remplie, il faut supprimer des éléments pour libérer de l'espace mémoire. Le numérique affiche alors un caractère nouveau, celui de pouvoir faire disparaître soudainement des contenus, dans cette même instantanéité qui nous permettait de les récupérer.

### 3.3.4 Machines en(chantées) vocales. *n-Cha(n)t*, de David Rokeby

#### *La voix comme instrument de contrôle*

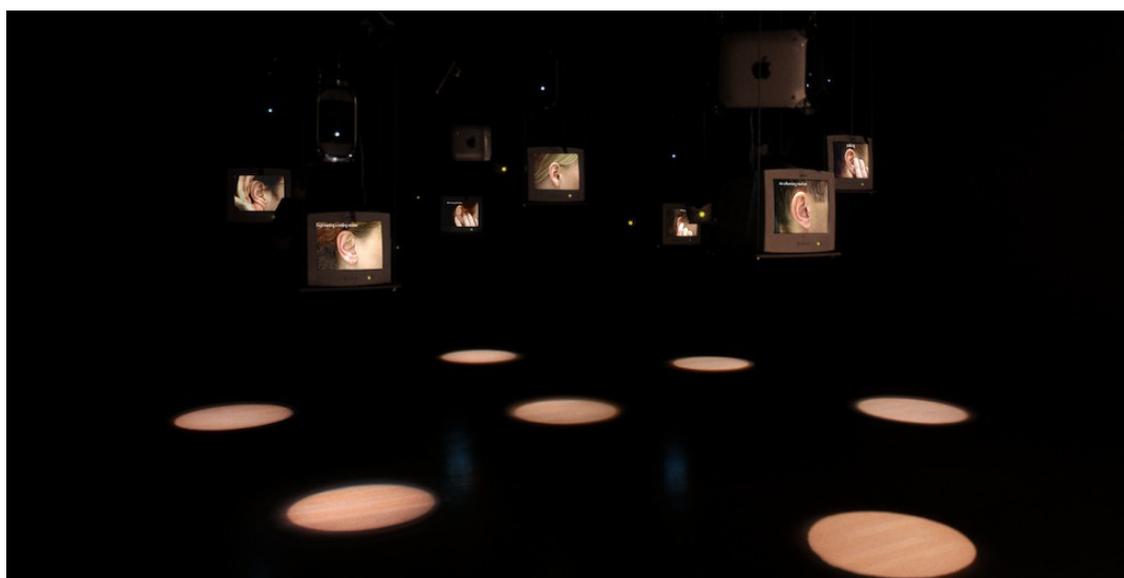


Fig. 178 : David Rokeby, *n-Cha(n)t* (2001)  
© David Rokeby / davidrokeby.com

Dans *Temps libre* de Samuel Bianchini<sup>674</sup>, l'interactivité mise en œuvre est relativement compréhensible par le visiteur : la voix du visiteur, selon son intensité, entraîne l'agitation visuelle de golfeurs. L'œuvre, par son dispositif direct, ouvre un questionnement sur l'addictivité au contrôle que nous pouvons développer dans le cadre d'un en-

674 Voir le premier chapitre de la deuxième partie, section 2.1.5.

vironnement dont nous comprenons progressivement, mais très rapidement les principes, à la manière d'un jeu vidéo.

Dans d'autres configurations, c'est la question du contrôle lui-même qui est en question. Que contrôlons-nous réellement ? Dans son installation interactive *n-Cha(n)t* (2001)<sup>675</sup>, l'artiste David Rokeby explore la question en proposant au visiteur d'interagir vocalement avec un réseau de machines informatiques parlantes.

### ***Des formes interactives de contrôle au dialogue avec les machines***

Dans un espace plongé dans la pénombre, le visiteur entend un ensemble de voix, tandis que sur une petite dizaine d'écrans sont diffusés des profils de visage. Chaque profil est rendu anonyme par un cadrage rapproché sur le pavillon auriculaire. Si le visiteur est seul, il est immergé dans une sorte d'unisson de voix synthétiques répétant toutes en boucle un même fragment de phrase énigmatique. Il pourra constater en se déplaçant qu'à chaque écran est associée une paire de haut-parleurs, qui diffuse l'une des voix répétitives. Sur chaque écran on peut voir la phrase prononcée par les haut-parleurs correspondants s'écrire mot à mot, en haut à gauche de l'écran, en synchronisation avec sa voix associée. Lorsqu'une phrase est prononcée, une main se pose sur l'oreille, comme un signe de protection du bruit extérieur. On assiste alors à un unisson machinique, composé d'entités possédant une incarnation visuelle (l'image et le texte) et sonore (la voix associée).

### ***La recherche de l'unisson***

Une fois accoutumé à l'ambiance de la pièce, le visiteur peut aussi remarquer la présence de micros sur chacun de ces emplacements ; ceux-ci lui indiquent qu'il peut avoir un rôle potentiel. En effet, si celui-ci émet un bruit, ou bien parle, les entités audio-visuelles les plus proches vont réagir.

---

675 Une captation vidéo est disponible en ligne sur le Vimeo de David Rokeby, <https://vimeo.com/17166940>, consulté le 30/01/23.



Fig. 179 : David Rokeby, *n-Cha(n)t* (2001)  
Walter Phillips Gallery, Banff Centre for the Arts  
© David Rokeby / www.davidrokeby.com

L'écran va montrer une main (cf. image gauche ci-dessus) qui se positionne en porte-voix derrière l'oreille, en signe de réceptivité, puis les haut-parleurs correspondants vont quant à eux commencer à diffuser des phrases, différentes de celles prononcées dans l'unisson. Ces nouvelles phrases, elles aussi, se voient affichées mot par mot. Cet acte dissident provoque une instabilité qui se propage et brise progressivement l'unisson.



Fig. 180 : David Rokeby, *n-Cha(n)t* (2001)  
Walter Phillips Gallery, Banff Centre for the Arts  
© David Rokeby / www.davidrokeby.com

Toutes les entités vont se mettre à recréer des phrases, créant cette fois une polyphonie qui semble anarchique. Si le silence est fait, le système va peu à peu regagner un état d'équilibre et revenir à un unisson, différent de l'état précédent. Si la réaction des entités est immédiate et directe, il est plus difficile d'appréhender son rôle en tant que visiteur.

Lorsque l'on parle devant le microphone, le système réécrit en bas de l'écran ce qu'il entend, et en extrait une liste de mots-clés. À partir de cette liste, et d'une base de connaissances préparée au préalable par l'artiste, le programme va fabriquer des phrases dont la grammaire est correcte, mais qui suivent des règles préétablies, avec une dose d'aléatoire.

Les ordinateurs sont connectés les uns aux autres et ils s'échangent entre eux leurs sujets d'intérêts du moment, et ainsi ils s'encouragent à penser aux mêmes choses. Et après quelques minutes, vous entendez que ce que cet ordinateur-ci raconte, ce que cet autre-là et celui-là au fond disent tournent autour de la même chose. L'un va commencer à parler de nourriture, l'autre dire qu'il a faim, puis parler de poulet, de poulet rôti, et tout doucement ils vont converger jusqu'à ce que soudain ils disent les mêmes mots, à peu près de la même manière, et vous allez vraiment vous croire dans une église remplie de gens qui prient<sup>676</sup>...

L'unisson est atteint car les ordinateurs s'entendent entre eux grâce à la reconnaissance vocale et finissent par utiliser progressivement le même registre de mots, ce qui aboutit à l'équilibre. Si l'un ne suit plus subitement le même registre, il instille un déséquilibre.

Il y a aussi dans *n-Cha(n)t*, la notion d'enchantement (*enchant*), que l'on peut associer à une certaine magie de la générativité. Le jeu entre les mots anglais *chant* (chant) et *chat* (discussion) questionne aussi le statut poétique du texte lu entre poésie et musicalité.

---

676 Retranscription de l'interview de David Rokeby par Xavier Ess pour l'émission Cybercafe21 (RTBF), au Théâtre Mercelis, Bruxelles, décembre 2003, <https://legacy.imal.org/drokeby/archives/interviewfr.html>, consulté le 30/01/23.

### **À notre écoute. La machine nous rassure**

Tout comme dans *Temps libre*, il y a une vie sans la sollicitation du visiteur. Dans les deux cas, le son provoque un dérèglement du système, et le silence fait revenir ce dernier à un point d'équilibre. La question dans *n-Cha(n)t* est de savoir si nous avons nourri le système. Nos mots ont-ils été intégrés dans la base, ou dans tous les cas y a-t-il leurre ? C'est-à-dire que nos mots donnés sont progressivement gommés pour revenir à la base de connaissance du dispositif. Le changement d'état de l'image — qui est une sorte de morphing — est à ce titre extrêmement pertinent. Il nous montre, par le geste du porte-voix, la machine qui prête l'oreille ; elle est à notre écoute.

Le dispositif semble si sophistiqué (les mots prononcés par le visiteur sont affichés au fur et à mesure, ce qui démontre le travail en tâche de fond de reconnaissance vocale), que l'on imagine que forcément, l'interaction est productive, enrichissante. L'image ici a pour rôle de nous mettre en confiance, tandis que tous nos mots sont analysés.

David Rokeby dit utiliser les technologies pour dénaturer des systèmes, qui ne nous comprennent pas et ne nous connaissent pas, dans l'idée de mieux nous examiner nous-mêmes<sup>677</sup>. Il cible en particulier l'invisibilité de technologies qui nous sont tellement familières qu'elles en deviennent dangereuses :

Je suis très intéressé par le problème que pose l'interface que l'on ne peut pas détecter et qu'on ne peut donc pas critiquer ni interroger. Toute interface de cette sorte crée inévitablement une distorsion de la perception de soi, puisque le comportement de l'interface invisible doit être incarné à l'intérieur de son propre modèle de soi. C'est la raison pour laquelle j'ai écrit au sujet des implications de la construction d'expériences par l'entremise de la programmation. Dans ce cas, le programmeur est en train de modifier les interacteurs, en changeant leur relation à eux-mêmes et avec le monde. Si cet acte de modification devient impossible à localiser ou à détecter, il devient alors politiquement problématique<sup>678</sup>.

677 Interview de David Rokeby par Melentie Pandilovski, directeur du Video Pool media Arts Centre (Winnipeg) en 2005, <https://www.videopool.org/watch-interview-with-david-rokeby/>, consulté le 30/01/23.

678 « David Rokeby : Je suis un artiste interactif et je construis des expériences », entretien avec Louise Boisclair, *Archée*, revue d'art en ligne : arts médiatiques & cyberculture, 2011, <http://arcee.qc.ca/ar.php?page=article&no=383>, consulté le 30/01/23.

Il est important de préciser que, tout comme dans la *Time Machine* de Piotr Kowalski<sup>679</sup>, les unités centrales et les claviers des ordinateurs utilisés sont suspendus et visibles dans l'installation, au-dessus ou avec les écrans (selon les modèles). Il y a en général 7 ordinateurs, mais le titre de l'œuvre, en utilisant le  $n$  mathématique, évoque une potentialité de fonctionnement avec plus d'ordinateurs.

Rokeby utilise dans *n-Cha(n)t* la base d'un programme génératif déjà utilisé dans son installation *The Giver of Names* (1990 —). Dans cette dernière, le visiteur pouvait placer sur un socle des objets choisis parmi ceux posés sur le sol. La forme, les contours et les couleurs de ceux-ci étaient reconnus par le programme, puis traduits en mots. Des phrases étaient ensuite générées et affichées par un moniteur.

### ***Écho face à Narcisse***

Dans bon nombre des installations dialogiques que nous avons étudiées, nous sommes renvoyés aux mythes d'Écho et de Narcisse. Écho avait été punie par Junon, fille de Saturne, à ne pouvoir répéter que les derniers sons émis par une voix.

Il advint que le jeune homme, séparé de la troupe de ses fidèles compagnons, cria : « Y a-t-il quelqu'un près de moi ? » « Moi », répondit Écho. Plein de stupeur, il promène de tous côtés ses regards. « Viens ! » crie-t-il à pleine voix ; à son appel elle répond par un appel. Il se tourne et, ne voyant personne : « Pourquoi, dit-il, me fuis-tu ? » Il recueille autant de paroles qu'il en a prononcées<sup>680</sup>.

La suite est connue, lorsqu'après ces jeux de reprise elle se montre, Narcisse la rejette. Écho, inconsolable, se voit alors dépérir, ses os devenir rocher ; elle ne se montre plus, « mais tout le monde l'entend ; un son, voilà tout ce qui survit en elle »<sup>681</sup>.

Face à la machine, nous tentons vainement de créer un lien, mais ne nous est renvoyée qu'une lecture de nous-mêmes.

679 Voir au premier chapitre de la deuxième partie, section 2.1.2.

680 Ovide, *Les Métamorphoses*, Livre III 383-410, Édition de Jean-Pierre Néraudeau, Paris, Gallimard, coll. « Folio Classique », 1992, p. 118.

681 *Ibid.* p. 119.

## 3.4 Talking Heads

### 3.4.1 Schèmes de l'extimité. Autour du travail de Natalie Bookchin

Nous avons, dans la deuxième partie, évoqué le travail de Natalie Bookchin avec les webcams<sup>682</sup>. Une grande partie de son travail est intrinsèquement liée à la culture web. D'abord protéiforme avec l'écllosion du net.art, l'esthétique de Bookchin se cristallise à la fin des années 2000 dans l'exploration de formes documentaires fondée sur le réemploi de *vlogs* (blogs vidéo sur le web).

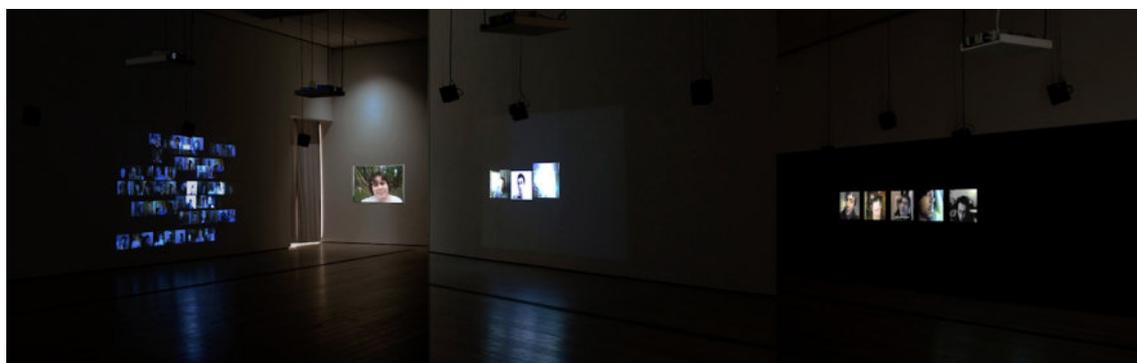


Fig. 181 : Natalie Bookchin, *Testament*, 2009/2017  
Vue de l'installation multicanal  
© Natalie Bookchin / bookchin.net

La série *Testament* (2008-2017)<sup>683</sup> est constituée de quatre vidéos linéaires, montrées ensemble dans une installation ou visibles séparément, sur son site web<sup>684</sup>, par exemple. Sur chacune d'entre elles, sur fond noir, nous observons et écoutons des personnes témoignant devant leur webcam d'une expérience personnelle sur une thématique que l'on comprend progressivement commune. Les témoignages sont directement adressés à la webcam, élargissant le jeu cinématographique du regard-caméra.

*Laid Off* (2009) traite du licenciement et du chômage, *My Meds* (2009) de la consommation de médicaments psychotropes, *I Am Not* (2009/2016) de l'identité sexuelle et

682 Voir le quatrième chapitre de la deuxième partie, section 2.4.1.

683 Des captations vidéo sont disponibles en ligne sur le Vimeo de Natalie Bookchin,

<https://vimeo.com/265980408>, et <https://vimeo.com/237790993>, consultés le 30/01/23.

684 Site Vimeo de Natalie Bookchin, <https://vimeo.com/user1977001>, consulté le 30/01/23.

*Count* (2008/2017) de la perte de poids. Pour chaque sujet, Bookchin a cherché, réuni, sélectionné et monté un ensemble d'extraits de *Vlogs*, publiés publiquement à différents moments et sur des sites divers.

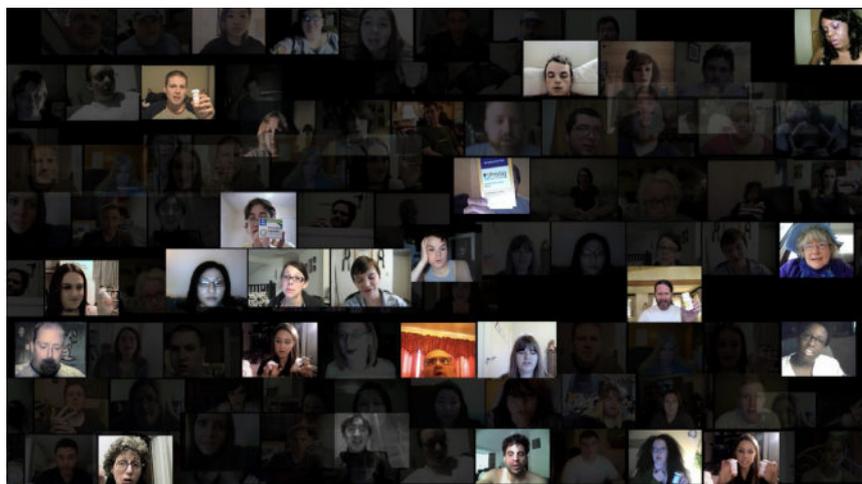


Fig. 182 : Natalie Bookchin, *My Meds* (2008)  
© Natalie Bookchin / bookchin.net

Le montage de ces témoignages vidéo, au format vignette, s'organise de manière singulière. D'une part spatialement, un grand nombre de vignettes vidéo – et donc autant de paroles – peuvent être présentes en même temps dans le cadre de l'écran, sur un modèle de *split-screen* en grille. D'autre part l'enchaînement des séquences recrée artificiellement des continuités entre les témoignages. Dans *Laid-Off* par exemple, la grille est seulement horizontale. La vidéo générale débute avec six séquences disposées horizontalement, les personnes prononçant toutes en même temps la phrase « So Today » [Et donc aujourd'hui]. Certaines séquences disparaissent, d'autres se déclenchent successivement ou simultanément, faisant naître une narration polyphonique. Quand des mots-clés (*Work, Time..*) [travail, temps] reviennent dans les témoignages, les séquences concernées se synchronisent et produisent des unissons de plus en plus intenses. Lorsque l'expression « Laid-Off » [renvoyé(e)], prononcée par tous les témoignant sur une ligne complète couvrant la largeur de l'écran retentit, l'effet est saisissant.

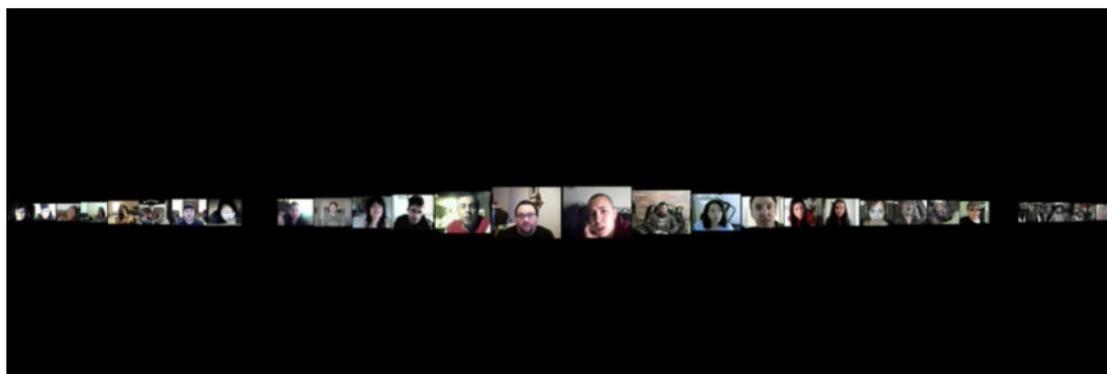


Fig. 183 : Natalie Bookchin, *Laid-Off* (2009), au moment de l'unisson  
 © Natalie Bookchin / bookchin.net

En effet, la plupart du temps, Bookchin compose l'espace de son écran avec un nombre réduit de témoignages individuels, sur un arrière-plan noir très présent. Ces vignettes, isolées, sont rejointes par d'autres lorsque les mêmes mots sont prononcés ; des multiplicités se créent, l'individualité fusionne avec le collectif.

### ***L'image fragmentée dans la fusion sonore***

Bookchin veille dans son montage à ce que les témoignages soient toujours audibles. Lorsque les quatre pièces vidéo sont installées ensemble, seul le son de celle qui est active est audible, tandis que les trois autres sont diffusées avec le son coupé.

Dans les moments d'unisson, qu'ils soient plus ou moins denses, le son qui était celui de la voix d'une seule personne, synchronisé à celui de son corps, semble se désincarner pour venir fusionner avec les autres ; une coalescence qui propulse le sonore au premier plan en mettant à distance le réel de l'individualité parlante. Tel un chœur, l'image fragmentée n'est plus que l'indice nous permettant d'évaluer le nombre de voix ; elle devient statistique.

Les similarités entre les témoignages ne sont pas que sémantiques, elles se manifestent aussi dans des gestes ou des expressions de visage similaires. Natalie Bookchin rappelle<sup>685</sup> à ce titre que même si ces témoignages restent poignants, le *vlog* est une

685 « Out in Public: Natalie Bookchin in Conversation with Blake Stimson », rhizome.org, mars 2011, <https://rhizome.org/editorial/2011/mar/09/out-public-natalie-bookchin-conversation-blake-sti/>, consulté le 30/01/23.

mise en scène de soi, la personne se regarde en miroir à l'écran, elle se cadre elle-même. Elle constate d'ailleurs un mimétisme avec des schèmes issus de la télévision, des manières de raconter ou de présenter les choses.

### ***L'extimité***

Serge Tisseron, dans son ouvrage *L'intimité surexposée*, définissait la notion d'extimité.

Je propose d'appeler « extimité » le mouvement qui pousse chacun à mettre en avant une partie de sa vie intime, autant physique que psychique. Ce mouvement consiste dans le désir de communiquer sur son monde intérieur. Mais ce mouvement serait incompréhensible s'il ne s'agissait que « d'exprimer ». Si les gens veulent extérioriser certains éléments de leur vie, c'est pour mieux se les approprier en les intériorisant sur un autre mode grâce aux échanges qu'ils suscitent avec leurs proches. L'expression du soi intime — que nous avons désigné sous le nom « d'extimité » — entre ainsi au service de la création d'une intimité plus riche<sup>686</sup>.

### ***Une esthétique numérique. Le réemploi et la requête***

La pratique du réemploi web chez Natalie Bookchin a débuté vers 2005, période à laquelle elle constate que le web est saturé et moins alternatif. De premiers travaux majoritairement *online* et interactifs, elle va passer à des vidéos *offline*. Le matériau est récupéré sur le web, mais les vidéos sont montées manuellement, avec des logiciels numériques standards. Elle utilise des méthodes de recherche (souvent sur Google) lui permettant de faire remonter des résultats moins populaires que ceux habituellement proposés, créant de cette manière un corpus plus personnel. Le caractère numérique de l'installation est ainsi forgé par le corpus vidéo de nature spécifique (le *vlog*), assemblé par une recherche maîtrisée dans les bases de données du web.

### ***Unissons et hoquets***

Dans les pièces de Bookchin se succèdent des moments d'unisson (une partie des voix prononcent le même mot en même temps), des séquences simultanées (dont certaines restent muettes), mais également des temps au sein desquels une phrase est fabriquée

---

686 Serge Tisseron, *L'intimité surexposée*, Paris, Hachette Littératures, 2002, p. 52.

artificiellement par un enchaînement de mots prononcés par différentes personnes. Cette distribution temporelle nous évoque très fortement la figure du Hoquet en musique, c'est-à-dire une « manière de faire alterner une ligne mélodique entre plusieurs voix<sup>687</sup> ». Jacques Siron, musicien et auteur, en propose la définition suivante.

Le hoquet est un procédé qui consiste à segmenter une ligne mélodique et à la distribuer entre plusieurs parties ; chaque partie ne joue que des petits fragments, parfois une seule note. Chaque partie est sans cesse en alternance avec les autres. Le hoquet se trouve aussi bien dans la musique médiévale (Machaut), dans la musique contemporaine (Webern) que dans certaines musiques traditionnelles (Afrique noire, gamelan indonésien). Dans le hoquet, l'effet d'enchevêtrement porte sur la distribution d'une ligne mélodique unique ; cependant il n'y a pas de superposition de plusieurs lignes indépendantes, à moins que la ligne fragmentée ne soit elle-même superposée à d'autres lignes. Le hoquet est une « polyphonie sociale » dans laquelle les parties sont fortement imbriquées les unes dans les autres : en effet, comme les échanges sont très rapides, il requiert une grande cohésion entre les interprètes<sup>688</sup>.

Appliqué à un ensemble de vidéos de parole, le hoquet serait donc la reconstruction d'une phrase par la succession des sons de vidéos isolées, l'image synchronisée au son ne devenant que secondaire dans la constitution d'un signifiant vocal général. La notion de polyphonie sociale, « microsociété dont les rapports sont constamment mis en jeu et remis en équilibre entre ses différents membres, chacun conservant sa propre autonomie<sup>689</sup> », que Siron développe dans son article s'applique tout à fait aux thématiques que Bookchin rend centrales.

---

687 « Mot du jour n° 102 : Hoquet », par Corinne Schneider, émission Le Mot du Jour, France Musique, Mardi 17 janvier 2017, <https://www.francemusique.fr/emissions/le-mot-du-jour/le-mot-du-jour-du-mardi-17-janvier-2017-31117>, consulté le 30/01/23.

688 Jacques Siron, « Enchevêtrement de voies », Cahiers d'ethnomusicologie, 1993, <http://journals.openedition.org.ezparse.univ-paris1.fr/ethnomusicologie/1431>, consulté le 30/01/23.

689 *Ibid.*

### 3.4.2 Candice Breitz. Théâtres noirs, une esthétique de l'isolation

Alors que Natalie Bookchin reconstruisait une narration documentaire à partir de récits d'anonymes s'adressant à la caméra, l'artiste Candice Breitz, dans son installation multicanal *Him + Her* (1968-2008), va au contraire détourner des séquences cinématographiques connues pour révéler les clichés de genre récurrents du cinéma hollywoodien.

#### *Un théâtre noir de portraits parlants*



Fig. 184 : Candice Breitz, *Her* (1978-2008)  
Durée : 23 minutes et 56 secondes  
© Goodman Gallery



Fig. 185 : Candice Breitz, *Him*, (1968-2008)  
Durée : 28 minutes et 49 secondes.  
© Goodman Gallery

L'installation vidéo bipartite nous présente un ensemble de fragments de scènes de l'actrice Meryl Streep d'une part<sup>690</sup>, un autre ensemble symétrique de scènes de l'acteur Jack Nicholson d'autre part<sup>691</sup>. Les arrière-plans des séquences vidéo d'origine ont été gommés par un détourage numérique pour laisser place à un fond totalement noir, à la manière d'un théâtre noir<sup>692</sup>. Seuls les visages des deux acteurs s'animent. Il s'agit à nouveau d'une choralité, fondée sur une polyphonie visuelle et vocale de portraits parlants.

Les deux ensembles sont composés de manière identique, sur la base de 7 écrans plasma : trois des écrans forment une colonne centrale, tandis que les deux autres paires constituent des colonnes centrées verticalement, à gauche et à droite. Cette disposi-

690 Montage d'extraits issus de 28 performances sélectionnées par Breitz sur 30 ans de carrière.

691 Scènes choisies parmi 23 performances sur 40 ans de carrière.

692 Dans ce type de théâtre, les acteurs vêtus de noir jouent devant un fond noir, ce qui donne l'illusion de visages parlants humains flottants.

tion permet à l'artiste de développer un langage narratif axé sur la composition spatiale. Elle peut créer des lignes, des colonnes, des obliques, des symétries, mais surtout, par contraste, des absences et des vides.

### ***Psychologies. La dépendance/L'égoïsme***

Chaque ensemble est dédié au même acteur évoluant dans différents rôles à différentes périodes. Chacun constitue un polylogue étrange de clones autocentrés. De plus, le rythme assez rapide des échanges verbaux (sortes de stichomythies<sup>693</sup> spatialisées), associé aux thématiques abordées, crée une tension permanente dans le déroulement narratif. La disposition des visages dans chaque ensemble donne la sensation qu'ils s'adressent les uns aux autres, dans un incessant passage de relais.

Au fil de l'expérience se révèlent dans ces deux ensembles des thématiques bien distinctes. Dans *Him*<sup>694</sup> [Lui], tout débute par une voix off « Tell us about yourself » [Parlez-moi de vous], puis les différents doubles de Nicholson enchaînent d'autres questions comme « Who I am? » [Qui je suis ?], et des réponses brèves « I'm a great guy » [Je suis un mec bien]. À travers son montage, Breitz dénonce un cinéma stéréotypé mettant en scène une image de la masculinité associée à l'autonomie et l'assurance. Au fur et à mesure des échanges verbaux, l'égoïsme masculin est progressivement mis à découvert, jusqu'à un point paroxystique : à dix-sept minutes et quinze secondes, on entend dans la bande-son un tir, suivi d'images de Nicholson à terre, en sang, pendant quelques secondes. Puis le ton change, des scènes de chant et de piano montrent une autre facette de l'acteur. Celui-ci semble avoir changé ; il présente une figure masculine délestée du poids de la représentation sociale lui incombant.

Pour *Her*<sup>695</sup> [Elle], aucune voix off. C'est le personnage d'une psychologue<sup>696</sup> qui entame le dialogue. « Tell me what it is, what you're trying to say » [Dites-moi ce qu'il y a, ce que vous essayez de dire], est sa première phrase, prononcée avec une légère émotion, adoptant le ton de celle qui en sait déjà long sur la réponse. Des vidéos bou-

693 Au théâtre, une succession rapide de très courtes répliques, au théâtre.

694 Une version est disponible en ligne sur le Vimeo de Candice Breitz, <https://vimeo.com/72198119>, consulté le 30/01/23.

695 Une version est disponible en ligne sur le Vimeo de Candice Breitz, <https://vimeo.com/72201223>, consulté le 30/01/23.

696 Il semblerait que la scène provienne de *Petites confidences (à ma psy)* (2005), réalisé par Ben Younger.

clés, comme dans *Him*, viennent ponctuer le déroulement de la narration. Certaines d'entre elles ne laissent entendre que des inspirations, ce qui crée une atmosphère anxieuse. Très vite, après un son d'alarme, les premiers thèmes énoncés sont le mariage, puis les tâches domestiques, les conseils maternels appuyés. Puis les entités de Meryl Streep commencent à parler, très émues, de leur relation de couple. Elles confessent leur dépendance affective et financière. À la moitié de la vidéo, certaines d'entre elles parlent d'écrire un livre, de trouver leur voie, les autres les encouragent ; cependant très vite le registre change : difficultés, échecs, mauvaise estime de soi. Très vite le mot divorce commence à circuler, « l'autre » a rencontré quelqu'un. À la fin de la vidéo, l'expression « to move on » [aller de l'avant], puis de la phrase « I'm taking back control on my life » [je reprends ma vie en main] viennent clôturer l'expérience dans un élan polyphonique optimiste.

### ***L'autonomisation des voix***

Souvent deux vidéos identiques, en miroir, se jouent simultanément. Dans le même temps, certaines séquences, de la mimique aux convulsions faciales, passent en boucle, silencieusement, dans un va-et-vient angoissant de spasmes qui rappellent les pièces structurelles de Martin Arnold, notamment *Passage à l'acte*<sup>697</sup>. La multiplicité des vidéos est parfois telle que l'œil n'est plus capable d'associer la voix à sa source, en particulier lorsque deux vidéos l'une à gauche, l'autre à droite, créent un unisson stéréo. Alors, comme dans *Testament* de Natalie Bookchin<sup>698</sup>, le son donne l'impression de se désolidariser de l'image. Le commissaire d'exposition Pascal Beausse, avait déjà relevé un effet similaire dans une œuvre plus ancienne de Breitz, *Babel Series* (1999).

L'effet de réel propre aux modes de narration audiovisuelle vise à créer une réalité spécifique, qui se pense autre que la réalité effective. Ce surplus de réel produit par l'espace cinématique est aboli par Candice Breitz. La présence des voix, qu'elle manipule comme des objets, échappe à leurs locuteurs. C'est une application non narrative du principe de la « voix acousmatique », décrit par Michel Chion : par la répétition radicale, la voix échappe en quelque sorte au corps qui l'émet, elle se décorpore. L'espace sonore créé par *Babel Series* s'autonomise des images, qui représentent des figures mécanisées par le

<sup>697</sup> Voir dans la deuxième partie, autour du saut et du va-et-vient temporel, section 2.2.3.

<sup>698</sup> Voir la section précédente, 3.4.1, autour de Natalie Bookchin.

rythme constant des boucles les enfermant. Les voix des pop stars sont bien émises par des effigies humaines, mais leur répétition et leur enchevêtrement créent un espace sonore spécifique, étrange, échappant aux corps. Des voix désincarnées, décorporées. Dans les mains de Candice Breitz, le montage est une arme de déconstruction<sup>699</sup>.

Même si plusieurs visages-voix cohabitent, le souci d'audibilité et de compréhension de ce qui est dit constitue un axe important chez Natalie Bookchin et chez Candice Breitz. Les moments d'unisson, contrôlés, correspondent à des accentuations de moments qui cristallisent une thématique particulière. L'image passe alors au second plan, et le sonore envahit la scène. Se produit un phénomène que l'on appelle en psychologie de la perception une capture sonore, c'est-à-dire une prédominance du sonore sur le visuel.

### ***Un dialogue de l'ombre double***

Avant ses mutations numériques, l'exercice de la parole filmée a connu un développement notable à partir des années 1960. D'une part, grâce aux caméras de cinéma portatives et leur capture de son direct, dans les formes journalistiques et documentaires ; d'autre part, grâce à l'utilisation des premières caméras vidéo et l'avènement de formes intimes et autocentrées (l'artiste se filme et se met en scène). Traditionnellement attribuée à Nam June Paik (il reste l'un des pionniers incontestables de l'art vidéo — notamment dans l'art de la distorsion du signal<sup>700</sup> —), la paternité de la première utilisation d'une caméra vidéo<sup>701</sup> semblerait pourtant revenir à Andy Warhol<sup>702</sup>, dans le contexte de son film *Outer and Inner Space* (1965)<sup>703</sup>. Le film, tourné en 16 mm, met en scène un monologue d'Edie Sedgwick<sup>704</sup>, assise à droite d'un moniteur qui diffuse la séquence sonore d'un autre monologue. Rien n'est numérique dans cette installation,

699 Pascal Beausse, « Candice Breitz : images infectées », texte de présentation de la soirée Candice Breitz, jeudi 28 avril 2005, Palais de Tokyo, <https://www.pointligneplan.com/document/candice-breitz-images-infectees/>, consulté le 30/01/23.

700 Notamment avec son exposition de 1963, *Exposition of Music – Electronic Television*, Galerie Parnass, Wuppertal, Allemagne. Commissaire : Rolf Jährling.

701 Une caméra Norelco de Philips.

702 Grégoire Quenault, « Histoires et mythes de l'art vidéo. Retour sur une genèse », in *Revue Ligeia* 2014/2 (N° 133-136), p. 161-170, <https://www.cairn-info.ezpaarse.univ-paris1.fr/revue-ligeia-2014-2-page-161.htm>, consulté le 30/01/23.

703 Une capture vidéo est disponible sur Media Art Net, <http://www.medienkunstnetz.de/works/outer-and-inner-space/video/1/>, consulté le 30/01/23.

704 Actrice et mannequin, devenue égérie de Warhol à l'époque de la Factory, avant sa mort en 1971.

mais sa configuration préfigure les expositions de soi contemporaines, du selfie aux réseaux sociaux.



Fig. 186 : Andy Warhol, *Outer and Inner Space* (1965)

Cette capture nous montre la version en diptyque. La face gauche montre la première bobine du film 16 mm, la face droite la seconde bobine

© The Andy Warhol Museum

Dans la vidéo du moniteur, filmée préalablement, Sedgwick est vue de profil, le visage tourné vers la droite ; elle semble parler avec naturel. Dans le film (le plan large de Sedgwick parlant devant le moniteur vidéo), elle s'adresse à une personne (plausiblement son retour vidéo) située hors champ sur la gauche et commente la situation dans laquelle elle se trouve : confrontée à sa propre image<sup>705</sup>. Ce dispositif donne l'impression qu'elle échange avec elle-même, dans une sorte de dialogue de l'ombre double<sup>706</sup>.

Au fil de sa parole, elle évoque des problèmes familiaux, la drogue, le décalage entre ses pensées (*inner space*) et la difficulté de les exprimer (*outer space*). Deux bobines d'environ 33 minutes ont été utilisées<sup>707</sup>. Dans sa diffusion originale, deux projecteurs diffusent les films côte à côte, en double écran ; la première bobine à gauche, la se-

705 « Andy Warhol, *Outer and Inner Space* », Media Art Net, <http://www.medienkunstnetz.de/works/outer-and-inner-space/video/1/>, consulté le 30/01/23.

706 *Dialogue de l'ombre double* (1985) est une œuvre de Pierre Boulez, dans laquelle l'instrument — la clarinette — livre un dialogue avec une bande préenregistrée. Le titre est une référence à une scène de la pièce de théâtre *Le Soulier de satin* (1929), de Paul Claudel, intitulée « L'ombre double », dans laquelle l'ombre est un personnage à part entière.

707 Callie Angell, « Doubling the Screen: Andy Warhol's *Outer and Inner Space* », *Millennium Film Journal* No. 38, 2002, <http://www.mfj-online.org/journalPages/MFJ38/angell.html>, consulté le 30/01/23.

conde à droite (voir l'image plus haut). Dans ce cinéma étendu<sup>708</sup>, chaque écran présente Edie Sedgwick et son « double » vidéo ; aussi nous voyons, et entendons, quatre manifestations simultanées de l'actrice car dans chaque bobine de film, se mêlent le son diégétique de la parole filmée et celui de la vidéo. Parfois Warhol effectue des zooms dans l'image et fait ajuster son volume en direct, mais l'ensemble présente néanmoins de nombreuses superpositions sonores qui empêchent régulièrement la parole d'être distinguée. Il est intéressant de noter que pour filmer en 16 mm, Warhol utilise une caméra en son direct<sup>709</sup>, appareil souvent associé au cinéma documentaire. Ceci explique le manque de relief du son, mais montre bien qu'il souhaitait recueillir une parole brute, tout en la mettant doublement en abyme par la superposition sonore et la spatialisation visuelle. L'avant-gardisme de cette œuvre étonne. Elle n'est pas numérique, mais, elle interroge déjà la parole filmée et ses multiplicités chorales, l'intime et la confession, le dialogue en miroir, la prise de son directe, la confrontation de techniques (la vidéo et le film), et enfin l'écran dans l'écran, le son dans le son.

### 3.4.3 Silences du regard

D'autres choralités contemporaines explorent le format du portrait vidéo, tout en se détachant de la parole filmée. L'installation *Terres arbitraires*<sup>710</sup> (2014), de l'artiste français Nicolas Clauss, présente un ensemble de moniteurs diffusant les portraits vidéo muets, ralentis, en noir et blanc, de jeunes hommes vivant dans les Zones Urbaines Sensibles (Z.U.S.) de villes françaises<sup>711</sup>. Les noms de ces quartiers (Les Avignonnets, la Z.U.P. de Surville...) sont, à différents moments, affichés simultanément sur tous les écrans. La bande sonore, spatialisée, nous donne à entendre ce qui semble être des captures in situ, ainsi que des sons de minuteur, de sirène, ou des extraits de reportage. Les jeunes sont filmés à l'extérieur, le cadre permettant d'entrevoir l'environnement quotidien dans lequel ils gravitent.

---

708 Nous faisons référence au terme introduit par Gene Youngblood dans son ouvrage *Expanded Cinema*, pour définir des formes cinématographiques audiovisuelles installées. Celles-ci peuvent être multi-écrans ou bien intégrer un système de diffusion sonore spatialisé (Gene Youngblood, *Expanded Cinema*, *op. cit.*).

709 Richard Ekstract, « La bande magnétique fait pop : un entretien underground avec Andy Warhol », *Tape Recording*, septembre-octobre 1965, in Andy Warhol, *Entretiens 1962-1987*, Paris, Grasset, 2005, p. 87.

710 Le titre fait référence à une citation d'Aimé Césaire.

711 Une captation vidéo est disponible en ligne sur le Vimeo du bureau Olivia S. Sappey d'Anjou, <https://vimeo.com/112859979>, consulté le 30/01/23.



Fig. 187 : Nicolas Clauss, *Terres arbitraires* (2014)  
© Nicolas Clauss / nicolasclauss.com

Le dispositif est modulable (selon l'espace d'exposition, différentes tailles et formats de moniteurs étant utilisés), mais il est surtout génératif. Le dispositif informatique en réseau<sup>712</sup> génère en effet des enchaînements semi-aléatoires de vidéos d'une part, de sons d'autre part. Il n'y a donc aucune synchronisation préalable, des associations entre images et sons se créent en permanence pour le spectateur-auditeur.

Quand j'ai commencé ce travail, j'ai fait un gros travail d'interview, notamment à Évry, et assez vite je me suis aperçu que ce serait peut-être plus juste de ne pas utiliser ce matériau sonore, d'utiliser vraiment le bruit médiatique, le bruit extérieur. Les gens qui sont filmés se mettent en scène, ils jouent pour la caméra : ils regardent la caméra donc ils regardent le spectateur, et c'est tout un jeu de regards, de dévisager l'autre, cette tension entre ces regards menaçants en début de séquence qui vont progressivement aller au ralenti, passer à des sourires, des fous rires. Il y avait donc un peu cette idée de confronter ces fous rires à des phrases abominables prononcées par les politiques<sup>713</sup>.

Nicolas Clauss avait déjà exploré ces portraits muets dans sa série de portraits vidéo-graphiques génératifs nommés *Endless Portraits*, qu'il débute en 2014. À la manière des

712 Basé sur le logiciel Max/MSP/Jitter.

713 Nicolas Clauss, Rencontre avec Xavier Thomas, Radio Grenouille, extrait, novembre 2011 retranscription, <http://www.nicolasclauss.com/terresarbitraires/TA.pdf>, consulté le 30/01/23.

*Screen Tests* (1963-1966) d'Andy Warhol (des films en plan rapproché poitrine de personnalités de la Factory, comme Lou Reed ou Edie Sedgwick), les personnes ne parlent pas, ce qui crée un malaise intéressant. Par ailleurs, grâce à la manipulation informatique par le code qui lui permet d'effectuer des sauts et des allers-retours temporels, un simple plan vidéo de 3 secondes peut devenir infini.

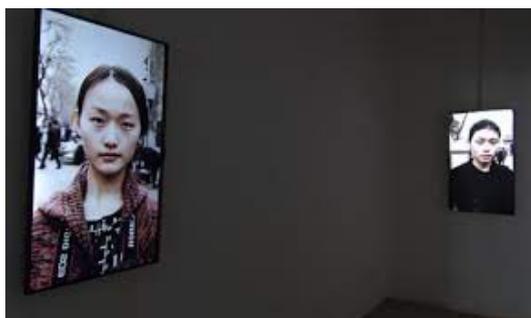


Fig. 188 : Nicolas Clauss, *Endless Portraits* (2015)  
© Nicolas Clauss / nicolasclauss.com



Il ne s'agit pas d'une boucle au sens interactif classique (dans ce cas la lecture de l'échantillon original se fait du début à la fin, puis revient au départ), ou d'étirement. Le procédé se rapproche plutôt de celui utilisé dans l'illusion auditive de la gamme de Shepard — sorte de chute sonore sans fin<sup>714</sup> — au sens où s'entremêlent des allers et des retours séquentiels. *Terres arbitraires*, comme les *Endless Portraits*, se déroule dans une temporalité particulière, elle n'a ni début ni fin.

<sup>714</sup> Nous avons évoqué ce paradoxe auditif avec l'œuvre *Colosseum* de Moritz Fehr, dans la deuxième partie, section 2.4.4.

### 3.4.4 *The Rumour of the World*. Une anthropomorphisation des *scams*



Fig. 189 : Joana Hadjithomas et Khalil Joreige  
*The Rumour of the World* (2014)  
 © Joana Hadjithomas et Khalil Joreige / hadjithomasjoreige.com

L'échange, le partage et la propagation d'emails constituent un noyau important dans la dynamique du web. Nous recevons régulièrement des spams, reçus parce que notre adresse mail a été retransmise ou revendue, mais également des cyberarnaques transmises par des personnes de notre réseau qui nous ont fait suivre un message qu'il pensait sincère. Cette forme de message, par lequel au terme d'un texte assez long une demande de don d'argent est faite, est appelée *scam*. Les artistes libanais Joana Hadjithomas et Khalil Joreige ont travaillé sur ces messages, plus particulièrement ceux basés sur une histoire personnalisée.

*The Rumour of the World*<sup>715</sup> (2014) est la deuxième partie d'un ensemble de pièces nommé *Scams (I must first apologise)*. L'installation<sup>716</sup> met en scène un ensemble d'écrans (plus d'une vingtaine) répartis dans un espace. Sur chaque écran figure une personne filmée semblant raconter une histoire, mais rien n'est directement audible car tout l'espace baigne dans une masse sonore constituée de ces différentes paroles. La sensation du visiteur est proche de celle d'une personne pénétrant dans un espace public peuplé, incapable de distinguer au premier abord quelle personne est la source de tel ou tel son. En s'approchant de l'un des écrans, on s'éloigne de l'agrégat central de

715 En français [La rumeur du monde].

716 Une captation vidéo est disponible en ligne sur le YouTube du Jeu de Paume, <https://www.youtube.com/watch?v=TTCJdfXk3ws>, consulté le 30/01/23.

haut-parleurs du plafond. Le son associé à un écran particulier va devenir plus audible — des haut-parleurs individuels sont placés près de l'écran, ce qui accentue l'effet — et va, par focalisation, accentuer l'effet d'adresse au visiteur. Celui-ci visiteur regarde et écoute alors le narrateur filmé se présenter, présenter des excuses préalables « I must first apologise » (« Je dois tout d'abord m'excuser... »), puis présenter le témoignage d'une histoire tragique. À un certain moment, il va s'agir de don, d'argent, il est alors clair que l'objectif de cette histoire est d'escroquer le destinataire. Les personnes filmées sont en réalité des acteurs qui lisent (ou plutôt interprètent) des *scams* récoltés et choisis par les artistes pour leurs « qualités » narratives.

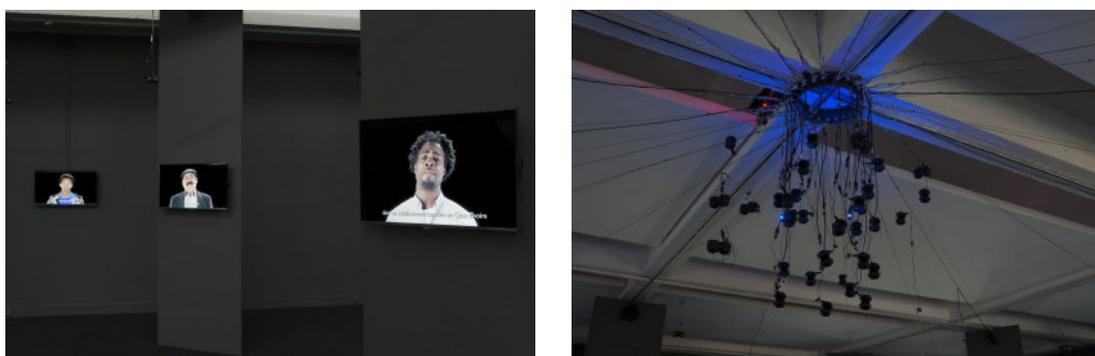


Fig. 190 : Joana Hadjithomas et Khalil Joreige  
*The Rumour of the World* (2014)  
 © Joana Hadjithomas et Khalil Joreige / hadjithomasjoreige.com

Le procédé de spatialisation sonore utilisé pour créer des focalisations n'est pas numérique, il s'agit d'un effet perceptif, mais il symbolise « l'accrochage » que peut constituer l'ouverture d'un mail contenant un *scam*.

Si l'approche du visiteur dans *Les Vagues* de Thierry Kuntzel<sup>717</sup> ralentissait peu à peu l'image jusqu'à la figer, ici c'est la présence du son qui se densifie pour recréer la synchronisation rassurante de la vidéo, par contraste avec l'ambiance sonore cacophonique dans laquelle débute l'expérience. Le continuum sonore dans lequel est plongé le visiteur symbolise le bruit numérique produit par la multiplicité des messages de ce type. Le visiteur approchant est le destinataire qui ouvre son email, pensant que celui-ci lui est destiné individuellement ; il est piégé et accroché. Dans l'installation plongée

<sup>717</sup> Voir dans le premier chapitre de la deuxième partie, section 2.1.4.

dans un brouillage sonore inconfortable, l'appât d'un son potentiellement plus habituel, intelligible, pousse le visiteur à aller s'approcher pour reposer son oreille<sup>718</sup>. Le musicologue et philosophe Peter Szendy remarquait que dans une installation, « l'artiste joue l'espace et laisse au spectateur-auditeur le choix de sa propre temporalité<sup>719</sup> ». L'artiste Max Neuhaus, pionnier dans l'utilisation du son dans les arts et premier à utiliser l'expression « installation sonore », en donnait cette description :

Les compositeurs situent habituellement les éléments d'une composition dans le temps. Une idée qui m'intéresse est celle de les situer plutôt dans l'espace, et de laisser l'auditeur les placer dans sa propre conception du temps. [...] Les installations sonores utilisent le son pour créer des lieux imaginaires d'exploration auditive, ou simplement pour se retrouver. Une œuvre se construit en choisissant les lieux sur la base de leurs réalités sonores dans un espace en composant un élément sonore chacun étant élaboré individuellement. Les combinaisons de sons qui proviennent de chaque source unique forment un ensemble qui se développe continuellement, à partir du moment où le visiteur entre dans l'installation, pour se conclure quand il s'en éloigne<sup>720</sup>.

Le dispositif met en scène la matérialisation d'une forme de message lié à la culture numérique (même si des formes de scam préexistent à l'ère numérique<sup>721</sup>). Il permet de créer des conditions d'écoute sélective, des appariements progressifs, tel un égaliseur permettant une focalisation sur la synchronisation labiale, de façon assez proche de l'effet *Cocktail Party*<sup>722</sup>. Dans l'installation *The Rumour of the World*, de nombreux haut-parleurs sont regroupés au plafond afin de créer l'ubiquité<sup>723</sup> initiale du son, c'est-à-

718 Une installation sonore de l'artiste Janet Cardiff, *Forty-Part Motet* (2001), pourrait être rapprochée, dans son principe de focalisation sur une partie du tout, de la pièce de Joana Hadjithomas et Khalil Joreige. Il s'agit d'une spatialisation de l'enregistrement d'une composition musicale polyphonique de Thomas Tallis, *Spem in Alium*, datant de 1573 (cette forme musicale est appelée un Motet). La particularité de l'enregistrement est qu'il est effectué voix à voix afin que chaque enceinte puisse diffuser isolément l'une des quarante voix qui composent la pièce originale. Le visiteur est incité à se déplacer au sein de l'agencement d'enceintes disposées en carré ou en cercle selon le lieu de présentation (qui peut être un espace muséal ou bien une église). Il peut ainsi changer de point de vue en s'approchant de chaque voix, contrairement à la disposition frontale classique des salles de concert.

719 Peter Szendy, « Installations sonores ? », in *Résonance* n° 12, septembre 1997, Ircam — Centre Georges-Pompidou.

720 Max Neuhaus, *Évoquer l'auditif*, catalogue d'exposition, Nice, Villa Arson, 1995, p. 101.

721 Joana Hadjithomas et Khalil Joreige, « La lettre de Jérusalem », in Joana Hadjithomas et Khalil Joreige, *Les rumeurs du monde. Repenser la confiance à l'âge de l'Internet*, Berlin, Sternberg Press, 2015, p. 49.

722 Colin Cherry, en 1953, nomme ainsi notre faculté à nous focaliser sur certains sons connus dans un contexte de bruit ambiant, tel un cocktail.

723 Petit lexique illustré des effets sonores, *op. cit.*

dire l'impossibilité que l'on a de le localiser ou de l'associer. Il n'est nulle part et partout à la fois.

Dans la relation affective qui se met assez vite en place dans les *scams*, l'adhésion à l'histoire ne relève pas simplement d'une naïveté, mais, pour reprendre les mots de Jacques Rancière, d'un régime de croyance à basse intensité<sup>724</sup>. Les artistes s'intéressent à ce moment où l'on décide de croire. Ces derniers ont aussi exploré l'autre face du miroir, celles des acteurs, non professionnels, en créant des pièces vidéo dans lesquelles ils donnent les raisons pour lesquelles ils ont souhaité incarner des textes de ce type. Ce sont, pour la plupart, aussi des histoires poignantes, ou décalées (l'un d'entre eux, Fidel, étant un ancien *scammer*). Ici le numérique n'est pas au cœur du dispositif technique, mais constitue le matériau de réflexion et de mise à distance.

---

724 « Jacques Rancière en conversation avec Joana Hadjithomas et Khalil Joreige », in Joana Hadjithomas et Khalil Joreige, *op. cit.*, p. 206.

## Quatri me partie.

### Carnet de recherche-cr ation. Une plate-forme exp rimentale web pour la recherche audiographique

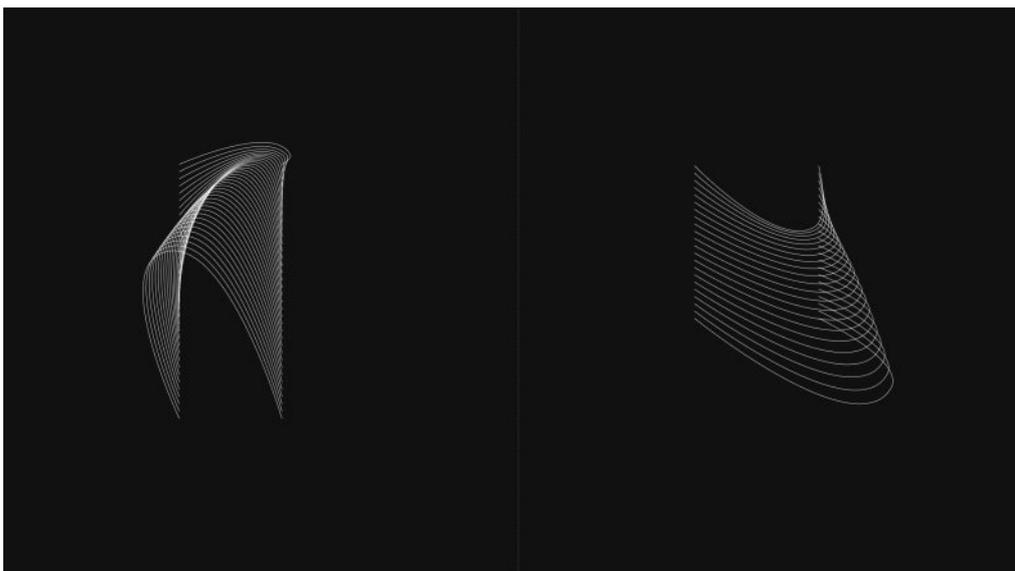


Fig. 191 : Capture de notre outil d'exp rimentation en ligne  
Deux entit s audiographiques  voluant en contrepoint sur deux parties d' cran  
  audiographic lab

L' criture de la premi re partie de notre  tude, « Formes de la musique visuelle programm e », a donn  conjointement lieu   une recherche artistique vou e    tudier les relations qui peuvent se cr er entre image anim e abstraite et son<sup>725</sup>. Celle-ci s'est mat rialis e dans l' laboration d'une plate-forme d'exp rimentation en ligne, *audiographic lab*<sup>726</sup>, que nous avons eu l'occasion de pr senter   diff rentes reprises<sup>727</sup>.

<sup>725</sup> Cette recherche, qui a d but    l'Universit  de Paris 8, est n e des croisements entre notre pratique de l'interactivit  et du code de cr ation, de notre enseignement en multim dia   l' cole nationale sup rieure des Arts D coratifs de Paris (EnsAD), et enfin de notre int r t pour les arts de l'image et du son. Toutes les exp riences d'enseignement et de recherche li es   cette th se sont donn es en annexe.

<sup>726</sup> Accessible   l'adresse <http://www.audiographiclab.com>, consult  le 30/01/2023.

<sup>727</sup> Le d tail de ces diff rentes communications est  galement donn  en annexe. Suite   ces communications, nous avons obtenu en 2017 une bourse du Fonds de soutien   l'initiative et   la recherche (Fsir) d'Arcadi, qui soutient « les initiatives d'exp rimentation et de recherche artistique et/ou culturelle, qui participent au renouvellement des formes et des esth tiques ». Cela nous a permis de d velopper le programme informatique et les interfaces n cessaires   nos premi res explorations. Gr ce   cet avancement, nous avons pr sent , en 2018 au Centre Pompidou, une version jeune public de notre plate-forme, pendant *La F te du code cr atif*. Cet  v nement  tait d di    la programmation informatique pour l'art. Nous d crivons ce travail dans l'une des derni res sections de cette partie.

La deuxi me et la troisi me partie, qui ont  t  ecrites ensuite, constituent les piliers pour de nouvelles extensions de cet outil, et donc de cette recherche-cr ation.

Nous pr senterons dans cette quatri me et derni re partie les principes sous-jacents et le fonctionnement de notre outil, nous d crirons des exemples de travaux, puis nous expliquerons de quelles fa ons le projet sera amen     voluer.

## **4.1 Des  tudes pour l'exploration de l'art « audiographique »**

### **4.1.1 Un art « audiographique »**

Dans la premi re partie, nous avons r uni des  uvres anim es abstraites, interactives ou g n ratives, au sein desquelles l'artiste tente d' tablir une relation d' quilibre entre la composante graphique et la composante sonore. Ni *musique* illustr e par de l'image, ni *animation* illustr e par de la musique, il s'agit pour les artistes   l'origine de ces cr ations d'explorer — pour reprendre les termes de John Whitney — une « musique audiovisuelle ». Comme nous l'avons d j  expliqu <sup>728</sup>, Whitney utilise ce terme, plut t que celui de « musique visuelle », pour poser une  quit  esth tique entre la composante visuelle et la composante sonore.

Dans le m me ordre d'id e, mais pour  viter l'utilisation du terme *musique*, trop connot , nous syst matiserons l'emploi d'un terme que nous avons parfois utilis  dans la premi re partie, celui d'*audiographique*. Il s'agit d'un terme existant<sup>729</sup>, mais peu r pandu<sup>730</sup>. Ce qualificatif induit d'embl e la pr sence de formes graphiques et de formes sonores, aussi il nous appara t tout   fait appropri  comme notion cl  pour caract riser les formes d velopp es dans le cadre de notre recherche.

### **4.1.2 D cloisonner les fondamentaux de l'animation et du sonore**

L'une des conclusions que nous tirons de notre g n alogie de la musique visuelle programm e — m me si elle n'est pas exhaustive —, c'est que hormis les quelques artistes qui ont tent  d' tablir des ponts interdisciplinaires (John Stehura, Lillian Schwartz,

---

728 Voir la section 1.1.4 sur les fr res Whitney, dans la premi re partie.

729 L'audiographie caract rise les repr sentations graphiques visant   pr senter les r sultats de tests d'audition ORL de patients, comme les audiogrammes.

730 Le terme est utilis  pour une qualification assez proche de la n tre dans le projet de recherche Topophonie (2009-2012), men  par l'enseignant et chercheur Roland Cahen.

John Whitney...), les champs de la musique informatique et ceux des arts visuels programm s sont rest s relativement cloisonn s jusqu'  la fin des ann es 1990. Puis,   cette  poque, les possibilit s technologiques sont si vite devenues  tendues qu'elles ont pris le pas sur la r flexion autour des relations audio-visuelles. Ainsi il est courant actuellement, dans des galeries ou des festivals d'art num rique, de voir des compositions graphiques et sonores g n ratives spectaculaires, mais qui n'ont, au fond, pas de structure : des feux d'artifice incontr l s. Les  uvres dont nous avons parl  dans la premi re partie ne sont pas de celles-l , mais elles sont finalement peu nombreuses. Selon nous, il ne s'est pas mis en place, comme pour le cin ma ou la musique, de grammaire ; ce qui pourrait s'apparenter   un « solf ge » audio-graphique. Pour le dire autrement, il y a encore peu d'endroits o  l'on apprend les basiques de la musique informatique, tout en apprenant les basiques du mouvement programm  : ces apprentissages restent souvent s par s. Partant de ce constat, nous avons postul  que le graphique et le sonore pouvaient se situer   un niveau d'attention  quivalent. Nous avons cherch    mettre en place des exp rimentations,   un niveau basique, de sons et de formes anim es  l mentaires, afin d'en faire  merger les points de convergence, les diff rences et les potentialit s.

Des *Five Film Exercices* des fr res Whitney dans les ann es 1940 jusqu'aux installations de Ryoji Ikeda dans les ann es 2000 qui mettent en sc ne des code-barres en noir et blanc, une grande partie des  uvres que nous avons  voqu es dans la premi re partie sont des  tudes, des exercices, ou bien encore des pi ces mettant en sc ne un vocabulaire graphique et sonore  l mentaire. La notion d'* tude* nous est alors apparue assez vite comme un passage oblig , un retour aux fondamentaux, avant d'aller vers des pi ces plus complexes. Les recueils d'«  tudes » (au sens de pi ces d'apprentissages) sont l gion dans la musique, mais il en existe  galement dans les arts graphiques. Il s'agit de pratiquer, chercher, t tonner sur la base d'exp rimentations, pour en  ventuellement d gager des figures : placer la pratique avant la th orie, comme le pr conise Josef Albers dans son ouvrage-manuel *L'interaction des couleurs*<sup>731</sup>. Il y explore la couleur par la pratique,   travers des exercices th matiques donn s : r manences, intervalles, illusions... Pour notre sujet en revanche, il nous semble que la

---

731 Josef Albers, *L'interaction des couleurs*, Paris, Hazan, 2013, p. 7.

couleur ne fait que multiplier les possibilit s, alors que nous pensons que nous devons nous focaliser sur la *relation* entre des formes anim es et des sons, qui  voluent dans un registre temporel. La couleur serait toutefois une strate d'exploration qui pourrait venir se greffer dans un second temps.

### 4.1.3 Le noir et blanc

Parmi les pi ces artistiques de la premi re partie, certaines nous int ressent pour leur utilisation de mat riaux  l mentaires. Nous avons constat  que la plupart  taient en noir et blanc, plus pr cis ment en blanc sur fond noir. Que ce soit chez des artistes-programmeurs tels que Manfred Mohr — qui a travaill  pendant plus de trente ans seulement en noir et blanc —, chez des artistes vid o tels que Woody et Steina Vasulka, ou bien encore chez des artistes contemporains comme Ryoji Ikeda, le noir et blanc est une contrainte que se donnent les artistes et qui favorise l' laboration d'une grammaire.

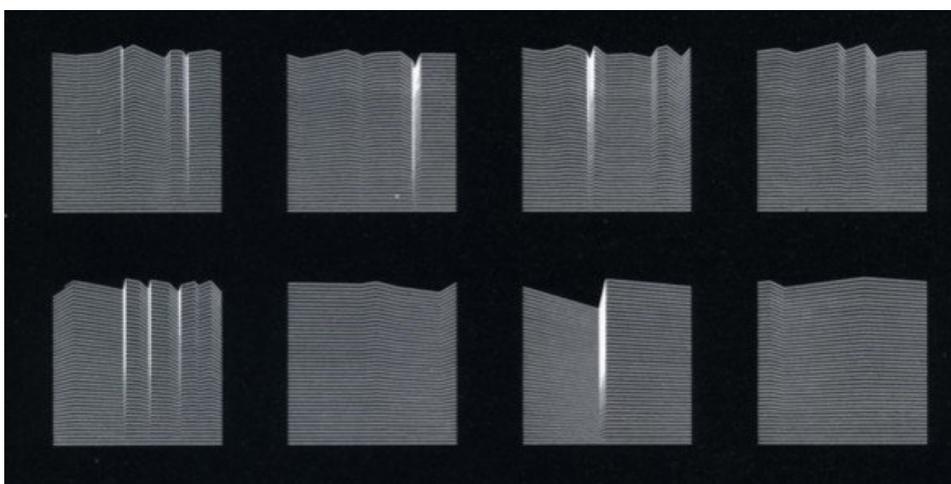


Fig. 192 : Manfred Mohr, *Matrix Elements* (1971)  
  ARC/Mus e d'Art Moderne de la Ville de Paris

Le caract re intemporel des premiers films abstraits de Viking Eggeling et de Hans Richter en atteste, le fond noir nous semble aussi appeler   la cin matique des formes, quand le blanc des figures favorise un contraste entre figure et fond propice   l'analyse. Dans le cas de pi ces audiovisuelles, l'attention port e au son semble meilleure quand la partie visuelle est  pur e et claire. D'une certaine mani re, et contrairement

  ce que pensaient les inventeurs d'orgues   couleur<sup>732</sup>, il nous semble plus int ressant de nous focaliser sur les relations de mouvement et d'interaction, que sur l' tablissement de correspondances de couleur. Les correspondances entre son et forme, quand elles se r alisent, sont plus g n ralement d'ordre temporel, comme l'imaginaient les fr eres Whitney.

#### 4.1.4 Les manuels d' tudes

Sur le mod le du Bauhaus, mais dans une perspective plus transversale et plus interactive,   l'image de laboratoires de recherche exp rimentaux tels que les Bell Labs ou le MIT (Massachusetts Institute of Technology) avec les travaux de John Maeda<sup>733</sup>, l'id e g n rale de notre projet est d'explorer simultan ment et de mani re interactive les basiques de la cr ation sonore (fr quence, timbre, balance, rythme...) et ceux de la cr ation graphique en mouvement (d placement, rotation, transition, vitesse...).

En compl mentarit  avec les pi ces artistiques d j   voqu es, nous nous inspirons de lectures crois es d'ouvrages et de textes traitant des basiques et des fondamentaux de diff rentes disciplines. Citons-en quelques-uns : Kandinsky qui d veloppe avec *Point et ligne sur plan*<sup>734</sup> une grammaire des formes, Theo van Doesburg et ses *Principes fondamentaux de l'art n o-plastique*<sup>735</sup>, Pierre Schaeffer et son *Trait  des objets musicaux*<sup>736</sup>, Miller Puckette avec *The Theory and Technique of Electronic Music*<sup>737</sup>, Gy rgy Kepes et sa s rie d'ouvrages *Language of Vision*<sup>738</sup>.

Dans le domaine du design graphique, l'un des ouvrages les plus int ressants reste pour nous le *Manuel de cr ation graphique*<sup>739</sup> du designer Armin Hofmann, qui propose une s rie d' tudes abstraites, toutes en noir et blanc, bas es sur des compositions de formes graphiques basiques comme le point ou la ligne. L'ensemble constitue une forme d'abr g  de son enseignement   l' cole de design de B le. Il y explore diff rentes polarit s graphiques li es   la trame, la densit , ou bien encore la distribution :

---

732 Voir la premi re partie.

733 Nous avons bri vement  voqu  le travail de John Maeda dans la premi re partie, section 1.3.4.

734 Kandinsky, *Point et Ligne sur plan*, Paris, Gallimard, coll. « Folio essais », 1991.

735 Theo van Doesburg, *Principes Fondamentaux de l'Art n o-Plastique*, Paris, ENSBA, 2003.

736 Pierre Schaeffer, *Trait  des objets musicaux*, Paris, Seuil, 1966.

737 Miller Puckette, *The Theory and Technique of Electronic Music*, Singapour, World Scientific Publishing Company, 2006.

738 Gy rgy Kepes, *Language of Vision*, Chicago, New York, Dover Publications, 1995.

739 Armin Hofmann, *Graphic Design Manual: Principles and Practice*, New York, Van Nostrand Reinhold, 1965.

« Le point dans un r seau de lignes », « Points de trois tailles diff rentes », « Jeu d'interf rences entre diff rents groupes de points », « R p tition des lignes verticales avec trois resserrements diff rents des espacements ».

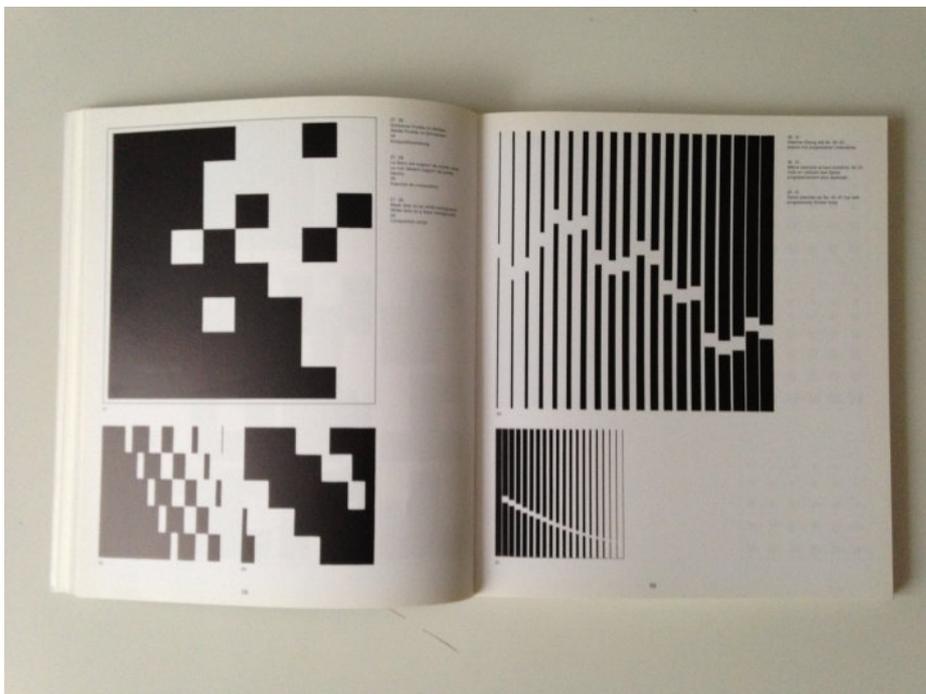


Fig. 193 : Armin Hofmann, *Manuel de design graphique*, 5<sup>e</sup>  dition, Niggli Verlag, 2004

Dans le domaine sonore, et au-del  des ouvrages de solf ge classique, il existe  galement des ouvrages similaires   celui d'Armin Hofmann, comme celui de John Pierce. Ce dernier  tait un ing nieur qui a travaill  sur la th orie de l'information, puis qui a appliqu  certains de ses principes   la musique. Il a notamment collabor  avec Max Mathews, l'ing nieur qui a donn  naissance   la musique informatique aux Bell Labs<sup>740</sup>. En 1983,   la suite de ses diverses exp riences, John Pierce  dite un ouvrage intitul  *Le son musical*<sup>741</sup>, dans lequel il d crit les aspects physiques et math matiques sous-jacents   notre perception de la musique.   cet ouvrage est joint un disque<sup>742</sup> sur lequel sont illustr s gr ce   la synth se sonore diff rents ph nom nes psychoacoustiques li s aux sons dits « musicaux ». Il s'agit d'une publication tout   la fois passionnante, pr -

740 Nous avons parl  de cette p riode dans la premi re partie, section 1.2.2.

741 John R. Pierce, *Le son musical. Musique, acoustique et informatique*, Paris, Belin, coll. « Pour la science », 1984.

742 D'abord un vinyle, puis un CD dans les r ditions.

cise et p dagogique. Il nous semble que ces «  tudes », pr sentes sur le disque, peuvent faire r f rence en termes de p dagogie d'initiation   la synth se sonore car elles sont encore valables aujourd'hui, quand bien m me les technologies ont  volu . Quarante quatre exemples y sont r partis en six parties, qui correspondent aux chapitres de l'ouvrage. Parmi eux, les basiques suivants peuvent constituer un ensemble d' tudes pouvant  tre mises en relation avec des composantes graphiques : « Ondes sinuso dales de deux fr quences diff rentes », « Battements », « Consonance et dissonance », « Construction d'une onde par ajout de partiels successifs », «  largissement et fusion de partiels », « Sons inharmoniques », « Transition entre deux timbres », « Canalisation des voix (Wessel) », « Son qui monte ind finiment »...

#### 4.1.5 Psychologie de la forme : vers une Gestalt audiographique

  partir du moment o  des formes multiples sont anim es dans un m me espace, se posent des questions de perception visuelle. Les lois de la Gestalt visuelle sont d sormais connues du grand public, notamment   travers les illusions optiques. Ce que l'on sait moins, c'est que ce type de recherche s'est  galement, au fil du 20  si cle, propag  au sonore.

L'origine de la psychologie de la forme remonte   la fin du 19  si cle, en Allemagne, plus pr cis ment en 1890 lorsque Christian von Ehrenfels introduit le terme « Gestalt » dans son ouvrage * ber Gestalt Qualit ten* (About Gestalt Qualities)<sup>743</sup>. Il y donne l'exemple de la m lodie, qui peut exister dans plusieurs tonalit s et   des rythmes diff rents. Au-del  d'une simple succession de sons isol s, elle poss de ses qualit s propres, qui persistent dans diff rents contextes. Dans le domaine visuel, les lois de la Gestalt visuelle sont  tablies par les psychologues Max Wertheimer, Wolfgang K hler, et Kurt Koffka, et  num r es par Wertheimer en 1923. La loi de proximit  d crit par exemple comment des fragments visuels rapproch s spatialement auront tendance    tre per us comme un tout. La loi de similarit  montre que des  l ments ayant une forme proche seront plus favorablement group s par notre perception. Dans le contexte d'un mouvement, la loi de destin commun  tablit que des composantes vi-

---

743 Jean-Marie Bolay, *Gy rgy Kepes. Du langage visuel   l'art environnemental*, Gen ve, M tisPresses, 2018, p. 30.

suelles ayant une trajectoire similaire viennent s'apparenter   un ensemble (les mouvements d'oiseaux en essaim par exemple).

En 1954, le th oricien de l'art Rudolf Arnheim sugg re que les principes d'organisation de la perception visuelle pourraient  tre appliqu s   la perception auditive<sup>744</sup>. Quelques dizaines d'ann es plus tard, l'analyse de sc nes auditives (ASA) d velopp e par Albert Bregman<sup>745</sup>, en constituera l'application fondatrice. Son ouvrage expose des recherches sur la mani re dont les sons sont fusionn s ou s par s par le cerveau dans la perception, certaines s'appuyant sur des transpositions dans le domaine sonore de lois de la Gestalt visuelle. Dans le contexte de ce type de recherche, on d couvre d'ailleurs des illusions auditives, moins connues que leurs cousines optiques (nous avons d j  parl  de la gamme infinie de Shepard, mais il en existe bien d'autres<sup>746</sup>).

Si leurs d veloppements respectifs ne cessent de progresser, la perception visuelle et la perception auditive sont souvent  tudi es de mani re cloisonn e. Autrement dit elles ne sont que rarement appr hend es de mani re *multimodale* ( tudi es simultan ment sur plusieurs sens). Nous savons pourtant que la vue et l'ou ie sont intimement li es, notamment dans la lecture labiale, mais aussi plus g n ralement dans les couplages entre  v nements visuels et  v nements sonores de la vie quotidienne. M me si nous n'avons pas ici tous les  l ments permettant d'explorer en profondeur ces aspects, il semble qu'une zone du cerveau comme le colliculus sup rieur soit le lieu d'un traitement multimodal audiovisuel<sup>747</sup>. Il existe d'ailleurs des illusions audio-visuelles, qui montrent qu'une approche multimodale peut s'av rer pertinente. L'effet McGurk, dont nous avons parl  dans la troisi me partie, met en sc ne un mot prononc  et un mouvement de l vre contradictoire. D'autres  tudes, fond es sur des formes gra-

---

744 Michael Haverkamp, *Synesthetic Design : Handbook for a Multi-Sensory Approach*, Berlin, Walter de Gruyter, 2012, p. 76.

745 Albert S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*, Cambridge, MIT Press, 1994.

746 Le chercheur Daniel Pressnitzer (co-directeur du Laboratoire des Syst mes Perceptifs de l' cole normale sup rieure et directeur de l' quipe Audition) a par exemple montr  qu'une s quence de notes diff rentes peut influencer la perception de hauteur de deux notes concluant une m lodie (Exp rience disponible en ligne, [http://audition.ens.fr/dp/illusion/index\\_fr.html](http://audition.ens.fr/dp/illusion/index_fr.html), consult  le 30/01/2023).

747 « Le colliculus sup rieur se distingue des autres centres sous-corticaux auditifs en ce qu'il est en m me temps une partie de la voie d'int gration visuelle et se r v le dot  d'une v ritable topographie de l'espace sonore [...]. Les unit s r pondant   l'audition r pondent  galement tr s souvent au stimulus visuel, de telle mani re que les champs visuels et acoustiques se r v lent en co ncidence spatiale. On peut donc dire que le colliculus sup rieur contient une repr sentation audiovisuelle de l'espace et que son fonctionnement implique de mani re pr f rentielle la pr sence simultan e des stimulus visuels et auditifs » (G rard Pel , * tudes sur la perception auditive*, Paris, L'Harmattan, coll. « Arts & Sciences de l'Art », 2016, p. 145-146).

phiques abstraites, mettent en lumi re des influences du son sur la perception visuelle ou leur r ciproque. Nous d taillerons l'une d'entre elles, le *Motion Bounce Effect* dans la section d di e   la description d'une s lection repr sentative de nos  tudes.

Cela  tant, sans applications concr tes   la cl , rares sont les scientifiques qui s'attaquent   ce sujet. Le cloisonnement artistique entre arts du son et arts du visuel que nous pointons dans notre travail existe aussi dans la science et la psychologie de la perception.

## **4.2 Un objet informatique audiographique : approche conceptuelle et technique**

### **4.2.1 La r alisation d' tudes en ligne**

Bon nombre des artistes dont nous avons  voqu  le travail (cette remarque est extensible aux deux autres parties) ont souvent une intuition, qu'ils cherchent ensuite   m rialiser en r alisant, ou en faisant r aliser des outils d di s. Les machines des fr res Whitney, le d veloppement en code de Stehura, le logiciel de John Whitney, les logiciels de Golan Levin, ou bien encore ceux d'Antoine Schmitt. De notre c t , nous travaillons dans le domaine du design web et du code de cr ation depuis le milieu des ann es 2000, et nous enseignons  galement ces pratiques et ces techniques en  cole d'art. Il nous a donc paru naturel de nous appuyer sur ces derni res pour r aliser notre propre couche de programmation d di e   l'exploration de l'art audiographique. M me en  tant   part enti re dans un processus num rique, l'id e est aussi que cela nous permette d' clairer des processus analogiques. En choisissant de d velopper nous-m mes l'outil, de le designer, de l'utiliser et de le rendre accessible, nous nous donnons  galement la libert  de le faire  voluer   notre guise.

Nous d crirons dans ce chapitre le programme qui permet de fabriquer des  tudes audiographiques, la description plus g n rale de la navigation dans notre plate-forme *audiographic lab* sera faite dans un autre chapitre.

## 4.2.2 Langages et techniques : HTML5, JavaScript et Web Audio

Les sites et les applications web modernes sont fond s techniquement sur ce que l'on nomme commun ment HTML5. En r alit , ce terme regroupe plusieurs technologies : HTML5, CSS3 et JavaScript<sup>748</sup>. Nous utilisons pour notre application cet ensemble, mais surtout, pour le c ur de notre programmation, le langage JavaScript. Il est l'un des langages de programmation le plus utilis  sur le web, et ce depuis les origines. Il permet d sormais la synth se et la manipulation de l'audio (Web audio), ce qui n' tait pas le cas auparavant<sup>749</sup>. L'avantage est que cela nous permet de cr er des programmes et des interfaces utilisables en ligne, en ayant la main sur la partie graphique des interfaces, et en travaillant de mani re « responsive ».

## 4.2.3 Des moteurs de synth se g n rative

Tout comme Ken Knowlton avait pr par  pour Lillian Schwartz le langage EXPLOR<sup>750</sup>, nous nous sommes cr   un programme d di    l'exploration de nos id es. Nous aurions pu repartir de z ro en programmant tout en JavaScript « pur », toutefois notre objectif n'est pas de tout r inventer, mais de croiser les pratiques. Il existe de nombreux modules (des *plugins*, des *frameworks*...), d j  programm s — c'est- -dire des couches orient es de simplification —, simplement ils sont souvent cantonn s   une nature de m dia, image ou son<sup>751</sup>. Nous avons donc d'abord choisi et r uni plusieurs moteurs de synth se existants, tous bas s sur le langage JavaScript, puis nous avons

---

748 HTML (HyperText Markup Language) est un langage de balisage qui ne permet qu'une chose : structurer s mantiquement un contenu. CSS (Cascading Style Sheets), quant   lui, est un langage servant   fabriquer des feuilles de style. Il permet de configurer toute la partie graphique de l'interface (typographie, couleurs, grille de composition...). JavaScript, enfin, est un v ritable langage de programmation, cr   en 1995, qui autorise la modification en temps r el de l'interface en fonction d' v nements d tect s (un clic de souris, un d filement de page...). Il permet non seulement la lecture de m dias (audio, vid o...), mais aussi la g n ration de textes, de formes graphiques...

749 Le son a longtemps  t  le parent pauvre du Web. Non seulement il fallait passer par des plugins tels que Shockwave (le plugin li  au logiciel Director) ou Flash, mais les possibilit s sonores se r sumaient principalement   la manipulation et au traitement de fragments audio. La g n ration synth tique sonore en temps r el  tait assez pauvre. Gr ce au Web Audio, de multiples projets en ligne ont fait leur apparition, comme le projet CoSiMa (Collaborative Situated Media) initi  par l'Ircam, qui proposait des s ries de performances audiovisuelles collectives   partir de smartphones (<https://cosima.ircam.fr/>, consult  le 30/01/23).

750 Voir dans la premi re partie, section 1.2.5.

751 Depuis le d but de notre travail, il faut toutefois noter que certains *frameworks* importants comme *p5.js* ont  volu  pour int grer de nombreuses possibilit s sonores.

cr e une biblioth que JavaScript<sup>752</sup> personnalis e, destin e   cr er des liens entre ces modules. Nous utilisons principalement trois moteurs pr existants.

Le premier moteur est charg  de la synth se sonore. Il doit pouvoir g n rer des sons de synth se, programmables et manipulables en temps r el. Nous utilisons *Tone.js*<sup>753</sup>, qui permet entre autres fonctionnalit s de changer en direct la tonalit , la spatialisation, l'enveloppe, gr ce   des combinaisons d'oscillateurs. C'est, en quelque sorte, une version simplifi e de programmes tels que *SuperCollider*<sup>754</sup>. Le deuxi me moteur doit permettre la synth se graphique, c'est- -dire, en quelques lignes de code, la g n ration de formes graphiques primaires (points, lignes, polygones...). Nous utilisons *Two.js*, une biblioth que capable de g n rer des formes au format SVG<sup>755</sup>, mais  galement susceptible de permettre des imbrications de formes anim es : des ensembles contenant d'autres formes,   des niveaux infinis. Par exemple, un disque qui se d place doucement et horizontalement de gauche   droite peut int grer un carr  qui rebondit rapidement, ce carr  contenant lui-m me un triangle clignotant (selon les  tudes nous utilisons  galement en alternance une autre biblioth que, *p5.js*, la version web de Processing<sup>756</sup>). Enfin, le troisi me moteur autorise la cr ation de courbes d'interpolation temporelle<sup>757</sup> qui peuvent s'appliquer aussi bien aux formes sonores qu'aux formes graphiques. Par exemple, si je cr e une courbe de transformation exponentielle sur une dur e fix e, je peux l'appliquer sur la tonalit  d'un son, mais aussi sur la position dans le temps d'un  l ment graphique. Si j'applique cette m me courbe   deux param tres – l'un sonore, l'autre graphique – alors que je r alise informatiquement une correspondance, autrement dit un couplage entre image et son, qui vont  voluer ensemble. Nous utilisons, pour la programmation de ces courbes, la biblioth que *GSAP.js* de GreenSock<sup>758</sup>. Elle permet  galement l'utilisation de timeline, c'est- -dire la cr ation

---

752 La traduction « librairie » pour « library » est souvent utilis e, mais le terme « biblioth que » est plus ad quat. Il s'agit de *plugins* pr programm s, directement utilisables, sortes de bo tes   outils document es. En combinant des modules, on peut cr er des programmes plus complexes ou puissants.

753 *Tone.js* est un framework Web Audio destin    cr er de la musique interactive dans le navigateur (<https://tonejs.github.io/>, consult  le 30/01/23).

754 *SuperCollider* est un logiciel de programmation et de synth se audio, cr e en 1996 par James McCartney (<https://supercollider.github.io/>, consult  le 30/01/23).

755 Le SVG (Scalable Vector Graphics) est un format standard de donn es capable de d crire des formes graphiques vectorielles.

756 Le framework *p5.js* est un d riv  du logiciel de code de cr ation Processing. Il permet de cr er des animations interactives et g n ratives publiables sur le web (<https://p5js.org/>, consult  le 30/01/23).

757 Par exemple les  quations de Robert Penner, qui d finissent diff rentes  volutions temporelles (<http://robertpenner.com/easing/>, consult  le 30/01/23).

758 Un *framework* d di    l'animation vectorielle (<https://greensock.com/gsap/>, consult  le 30/01/23).

de partitions temporelles. Elle peut potentiellement permettre d'animer n'importe quelle propri t  graphique ou sonore. Chacune de ces biblioth ques peut, sous certaines conditions,  tre remplac e par une autre, selon les besoins mais de mani re transparente. Elles sont interchangeables. Ce qui a permis   des d veloppeurs de cr er ces moteurs, c'est la capacit  de JavaScript    tre utilis  en mode « objet ». C'est- -dire, pour simplifier, de pouvoir ajouter ses propres objets, qui ont leurs propres comportements. C'est  galement ce que nous faisons, avec notre programme qui permet de cr er, par le code, des couplages entre des param tres<sup>759</sup>. Voici un sch ma qui synth tise le processus de fabrication d'une « animation audiographique<sup>760</sup> ».

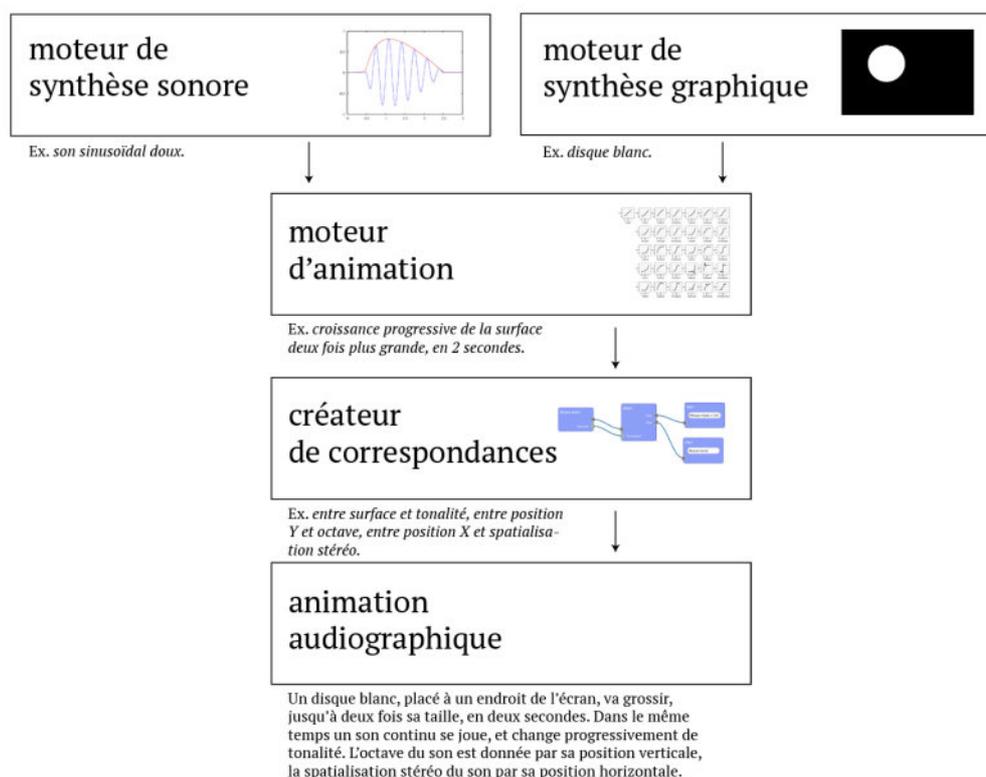


Fig. 194 : Sch ma de synth se du processus de fabrication d'une animation audiographique

Plus encore, le JavaScript, en tant que langage d'interaction, permet directement la d tection de clics, de survols, de d filements d' cran, d'appuis sur des touches du cla-

<sup>759</sup> Bien s r des logiciels de programmation visuelle tels que *Max/Msp/Jitter* auraient pu  tre utilis s, mais nous avons choisi une architecture qui nous permette de travailler facilement la partie graphique de l'interface et pouvant  tre mise en ligne.

<sup>760</sup> Le cin aste d'animation Norman McLaren, qui g n rait des sons en dessinant directement sur la pellicule, ou bien par le biais de cartes, parlait de « son anim  ».

vier. Chaque  tude nous permet ainsi d'associer des couplages entre param tres sonores et graphiques, mais aussi entre gestes interactifs et g n ration audiographique. Nous pouvons  galement, en temps r el, changer ou faire  voluer n'importe quel param tre graphique ou sonore.

#### 4.2.4 Configuration de l'entit  audiographique

Suite   de multiples exp rimentations avec les moteurs de synth se graphique et sonore, nous avons d fini un certain nombre de formes et de sons de base, suffisamment repr sentatifs pour qu'ils se distinguent lorsqu'ils sont jou s ou anim s ensemble. Nous avons donc   notre disposition une biblioth que de formes et de sons. Concernant ces derniers sons, il peut s'agir de sons doux, m talliques, percussifs, ou bien de bruits blancs. Il existe diff rentes mani res de les g n rer et de les fa onner, comme le travail sur l'enveloppe dite ADSR<sup>761</sup> du son.   la mani re d'une touche de synth tiseur, sur laquelle on appuie longtemps, il est possible de programmer le son pour qu'il se maintienne pendant une dur e donn e. Cette biblioth que est ouverte et sera r guli re nourrie avec de nouvelles configurations sonores et graphiques.

Sur la base de ces biblioth ques, nous avons conceptualis  informatiquement un objet audiographique, entit  primaire, r ceptacle, lieu de croisement, susceptible d'accueillir une synth se g n rative de formes graphiques anim es et de sons. Nous en avons d fini les r gles et les comportements.

Un objet audiographique peut  tre   la fois graphique et sonore. Dans ce cas, nous mettons en correspondance (une synchronisation) une animation de forme avec un son, qui se d veloppent dans un d veloppement temporel commun. Par exemple, pendant 5 secondes, un point blanc, sur un fond noir, se d place de gauche   droite de l' cran, tandis qu'un son sinuso dal continu,   une note donn e (un do m dian par exemple) se joue sur le m me temps en passant de l'enceinte gauche   la droite. Il y a alors deux correspondances, spatiales et temporelles. Je peux aussi choisir de lancer cette animation seulement au clic. Si je clique   nouveau sur l' cran, cela en g n re une nouvelle, avec  ventuellement des param trages diff rents. L'objet audiogra-

---

<sup>761</sup> D signe la forme du signal d'amplitude du son (son profil), mod lis e par le biais de quatre param tres principaux : l'attaque, le d clin, le maintien, le rel chement (Attack, Decay, Sustain, Release). En modifiant ces param tres (par le code, ou bien en manipulant directement la courbe graphique dans une interface), il est possible de g n rer une multitude de natures de sons.

phique peut aussi  tre seulement sonore, ou seulement graphique. Nous pouvons ainsi ind pendamment jouer des sons et animer des formes sans couplage pr alable, et tester des situations de d synchronisation.

Nos animations audiographiques sont techniquement fond es sur une entit  informatique abstraite. Celle-ci est charg e, pour le dire simplement, de cr er des relations entre un timbre de son, une forme graphique, et une nature d'animation.

La question de l'espace de l' cran est importante. Pour en simplifier l'occupation, nous avons cr e un syst me de grille (une sorte de quadrillage). Nous cr ons tout d'abord la zone de visualisation de la composition (la sc ne ou stage en anglais), puis nous la divisons en lignes et colonnes (autant que souhait ). Nous donnons ensuite   l'animation des coordonn es de d part et d'arriv e dans le contexte de cette grille. Si nous changeons en direct les dimensions de la sc ne (pour composer, tester, mettre en plein  cran, utiliser un t l phone portable), alors toutes les animations vont se recomposer proportionnellement (ou selon d'autres r gles).

```
var ag2 = new AudiographicObject({
  stage:stage2,
  group:"deplace",
  durationMS:2000,
  repeatDelayMS:1000,
  formName:'disque',
  w:10,
  h:10,
  fill:true,
  synthName:'SimpleSynthB',
  isPerc:false,
  params:
  [
    {type:'graphic',name:'left',minValue:1,maxValue:7},
    {type:'graphic',name:'top',minValue:3},
    {type:'audio',name:'frequency',minValue:100},
    {type:'audio',name:'pan',minValue:-1,maxValue:1}
  ]
});
```

Fig. 195 : Exemple simple d'un code de g n ration d'une entit  audiographique

Nous traduisons le code de l'exemple ci-dessus dans l'ordre des lignes d'instruction. Une animation nomm e *ag2*, et plac e sur la sc ne *stage2* (une grille de 7x5 configur e   un autre endroit du programme) de 2000 ms (2 secondes) est cr e. Elle se r p tera apr s 1000 ms (1 seconde). Il s'agit d'un cercle blanc (le fond est toujours noir et les formes blanches) de taille 10x10, rempli (un disque), le son pr d fini est le son

*SimpleSynthB*, qui n'est pas un son percussif, et sera donc jou  en continu toute la dur e de l'animation. La partie « params » concerne les correspondances. Nous indiquons   la forme graphique (pendant la dur e de 2 secondes) de se d placer de la colonne 1   la colonne 7 (de gauche   droite),   une hauteur de 3 lignes, avec un son jou    une fr quence fixe de 100 Hz, et passant de l'enceinte gauche   l'enceinte droite (le pan d signe la spatialisation st reo, c'est-dire la balance).

Il s'agit d'une des configurations les plus simples, mais elle montre comment nous pouvons activer des correspondances entre des param tres graphiques et sonores. Il faut toutefois comprendre que la configuration de ces couplages n'est pas une  tape obligatoire. Nous pouvons g n rer d'un c t  un ensemble d'animations audiographiques muettes, de l'autre un ensemble de sons sans visuel.

L'exemple du disque blanc sonore et anim   voqu  pose d j  question. Un musicien y verra peut- tre une forme graphique guid e par un son, un animateur y entendra la sonorisation d'un mouvement. Le principal objectif de l'application est de pouvoir ainsi tout aussi bien manipuler ind pendamment des formes sonores, des formes graphiques, de cr er des ancrages artificiels (du mappage entre image et son) audio-graphiques, ou enfin des sortes de contrepoints. Penser le son et le mouvement graphique d s la naissance de l' tude.

### **4.3 Quelques exemples d' tudes**

Sur la base des inspirations et mod les d crits dans les sections pr c dentes, nous avons imagin  une s rie d' tudes interactives   l' cran dans lesquelles se croisent les fondamentaux de la cr ation sonore (variations d'intensit , de fr quence, de composition spectrale, jeux de partiels<sup>762</sup>...) et ceux de la cr ation graphique anim e (variations de forme, de position, de rotation, d' chelle...). Ces  tudes sont organis es par th matiques (le bruit, l' cho, la vibration...). Nous partons de situations  l mentaires : « Un point anim  et un son », « Deux lignes droites anim es et deux sons ». Nous  toffons ensuite ces explorations, jusqu'  travailler sur des flux d' l ments graphiques et de grains sonores r gis par des algorithmes plus complexes.

---

<sup>762</sup> En sciences de l'acoustique, on dit qu'un son musical est d composable en composantes  l mentaires, qui se situent chacune   une fr quence particuli re : les partiels.

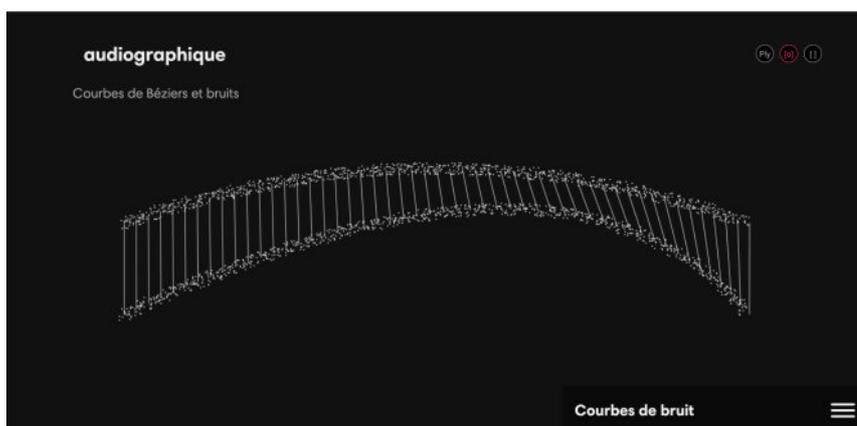


Fig. 196 : Forme composite en mouvement  
  audiographic lab



Fig. 197 : Courbe de B zier et bruit  
  audiographic lab



Fig. 198 : « Rubans » tournants  
  audiographic lab

Plusieurs m thodes coexistent pour notre recherche. Comme ce type d' tude n'existe a priori pas (si ce ne sont les  uvres  tudi es), nous travaillons   partir d'id es personnelles ou bien de ressources existantes. Nous nous inspirons aussi de formes trait es dans une modalit  seule (sonore, visuelle, interactive, algorithmique...), pour imaginer comment pourraient se mat rialiser les modalit s « manquantes ». Nous d crivons dans ce chapitre diff rents processus de fabrication de ces  tudes.

### 4.3.1 Une  tude fondamentale : « un point anim , un son »

#### *Le point dans son plan*

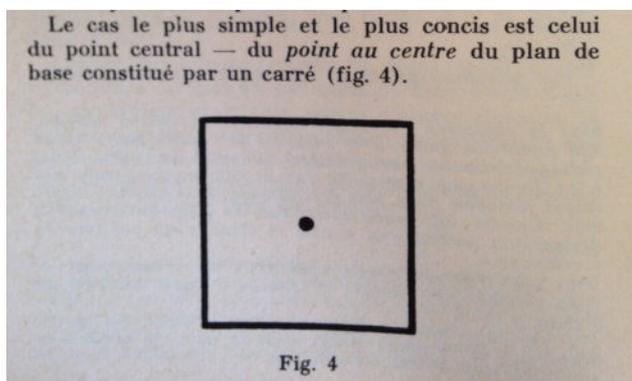


Fig. 199 : Vassily Kandinsky, *Point et Ligne sur plan* (1926)

Faire se croiser des fondamentaux de l'animation graphique et du son induit la mise en place de situations qui peuvent para tre simplistes. Il nous apparaissait pourtant primordial de d biter notre s rie d' tudes par des interactions entre un point blanc (sur fond noir, comme toutes les  tudes) et un son. Cela a, en effet, le m rite de poser des questions essentielles : qu'est-ce qu'un point ? Qu'est-ce qu'un son ?

En 1926, dans un chapitre d di  de son ouvrage *Point et ligne sur plan*<sup>763</sup>, Kandinsky d crit la symbolique du point, en partant notamment de son caract re intrins quement silencieux dans le contexte de la typographie. Il d crit ensuite les tensions qui peuvent  merger selon la position d'un point abstrait dans le plan, puis la complexit  des forces  manant de la mise en plusieurs points, et la r solution   apporter dans la composition pour g rer ces instabilit s potentielles.

Les *dimensions* et les formes du point varient et de ce fait change aussi la r sonance relative du point abstrait. *Ext rieurement* nous pouvons d finir le point comme la plus petite forme de base. Pourtant, cette d finition n'est pas pr cise. Il est difficile de d finir les limites de la notion « la plus petite forme » — le point peut grandir, devenir surface et remplir imperceptiblement toute la surface de base, — o  serait alors la limite entre point et surface<sup>764</sup> ?

<sup>763</sup> Vassily Kandinsky, *Point et Ligne sur plan*, Paris, Gallimard, coll. « Folio essais », 1991.

<sup>764</sup> *Ibid.*, p. 37

Se pose donc en premier lieu la question de la taille du point dans le plan. La relation entre figure et fond est fondamentale dans les arts graphiques. Pour Kandinsky il est n cessaire de tenir compte de deux conditions particuli res : d'une part le rapport entre les dimensions du point et celles de la surface de base, d'autre part le rapport de ces dimensions avec les autres formes sur cette m me surface. Il exprime donc les dimensions « id ales » d'un point de mani re relative —   la surface — et non pas de mani re absolue. Le peintre travaillant directement sur la surface finale, il est   m me d'appr cier les  chelles ; mais selon que le format soit miniature ou tr s large, la taille du point ne sera  videmment pas la m me. Nous nous situons d s lors dans des probl matiques d' chelle, questions auxquelles le graphiste est constamment confront . Un document ou une cr ation graphique compos s sur un certain type d' cran ne donnera pas le m me r sultat sur un autre appareil, ou lors d'une projection. Ce qui nous semblait  tre un point devient un disque, ou ce qui nous semblait  tre un disque devient un point. Dans le cas du point statique, Kandinsky effectue m me une analogie avec la musique :

La stabilit  du point, son refus de se mouvoir sur le plan ou au-del  du plan, r duisent au minimum le temps n cessaire   sa perception, de sorte que l' l ment temps est presque exclu du point, ce qui le rend, dans certains cas, indispensable   la composition. Il correspond   la br ve percussion du tambour ou du triangle dans la musique, aux coups secs du pivert dans la nature<sup>765</sup>.

Dans notre cas, nous nous situons dans un contexte de composition de formes graphique anim es et de fragments sonores, la variation dans le temps est incontestablement un  l ment de structure commun aux deux natures de composantes. L'analogie intuitive entre un son percussif,  ph m re, et un point nous para t en ce sens tout   fait int ressante   explorer *in situ*. D s lors que le point est potentiellement « animable », nous pouvons  tablir des couplages entre diff rents param tres graphiques et sonores.

---

765 Vassily Kandinsky, *op. cit.*, p. 42.

## Un point, un son

Ces pistes d'exploration nous permettent d'envisager une premi re s rie d' tudes, autour du point fix  graphiquement   un endroit donn  (statique), coupl  avec des sons simples, de nature percussive ou bien continue.

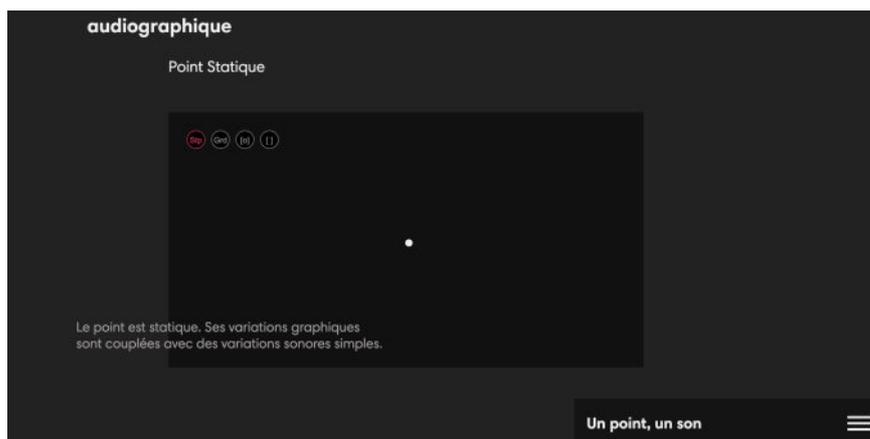


Fig. 200 : La surface d'un point statique, centr , cro t, tandis que se joue un son continu  
  audiographic lab

Une  tude va mettre en sc ne un petit disque blanc centr  sur l' cran (cf. ci-dessus). Tandis que sa surface cro t jusqu'  une certaine  chelle (un seuil que nous avons fix ), un son continu se joue,   une fr quence donn e, son volume augmentant en fonction de la surface. Nous pouvons aussi reproduire l'exp rience avec un son percussif.



Fig. 201 : Un point se d place de gauche   droite de l' cran, un son se joue et passe de l'enceinte gauche   droite  
  audiographic lab

Une autre  tude (cf. ci-dessus) pr sente un point qui se d place de gauche   droite, accompagn  d'un son continu. Dans ce cas pr cis, on remarque que si le son n'est pas spatialis , le son semble associ    l'animation g n rale, tel un arri re-plan.

Si l'on fait passer le son de l'enceinte gauche   l'enceinte droite, alors le son semble synchronis  — comme ce serait le cas pour un moustique —   la forme en mouvement : il y a une sorte de similarit  de trajectoire. Toutes les analyses qui d coulent de ces «  coutes-observations » seront int gr es sous forme de textes dans la plate-forme.

### ***Points multiples, son multiples***

Le designer graphique Armin Hofmann, dans son *Manuel de design graphique*, d die comme Kandinsky un chapitre au point. Il y d crit la situation dans laquelle est ajout  un second point sur la surface, puis d'autres.

Si nous placons un second point   c t  du premier, le rapport point/surface, qui jusqu'ici  tait le seul existant, devient secondaire. Ce sont deux points qui d cident maintenant de ce qui se passe sur la surface. Chaque point dirige ses forces vers l'autre en trac  lin aire. Si on les dispose convenablement, ils peuvent couper la surface en deux parties, sauter en dehors du format. Si nous diminuons la distance entre eux de telle sorte qu'ils se rencontrent, le couple de points se cr e, qui, dans une fusion croissante, engendre de nouvelles formations de points les plus diverses. Les rapports lin aires qui naissent de trois points dispos s en triangle cr ent un circuit ferm  de forces. Les  changes se font   l'int rieur du format. Un grand nombre de points donne un grand choix de formules : simple alignement de points, alignement vertical/horizontal (formation de trame), groupement, dispersion libre ou ordonn e, concentration, vari t  de dimensions, de gris, de couleur, de structure<sup>766</sup>.

Toutes ces possibilit s d crites par Hofmann, li es   la mise en sc ne de points multiples, nous inspirent bien entendu des transpositions audiographiques. Nous avons ainsi commenc    cr er des  tudes de points multiples.

---

<sup>766</sup> Armin Hofmann, *op. cit.*, p. 27.



Fig. 202 : Dans l' tude ci-dessus, tous les points se d placent de gauche   droite, sur une m me trajectoire horizontale,   des vitesses al atoires. Dans une autre, tous les points partent du coin gauche en haut et se dirigent vers le coin droite en bas. Chaque mouvement de point est mis en correspondance avec un son continu spatialis  de gauche   droite.

  audiographic lab

Kandinsky ou Hofmann, dans leurs textes respectifs, d veloppent  galement leurs r flexions autour d'autres formes  l mentaires, comme la ligne. Nous donnons donc ici des exemples de la mani re dont nous pouvons partir de textes, tir s d'ouvrages d'esth tique ou de design pour imaginer des  tudes. Dans de futures  tudes, nous partirons, sur un principe similaire, de situations sonores d crites dans des ouvrages d di s au son, telles que celles d crites dans l'ouvrage de John Pierce, *Le son musical*, que nous avons  voqu  dans le premier chapitre.

### 4.3.2 Isolation et reprogrammation de micro-s quences : exemples avec Oskar Fischinger et John Whitney

#### *Des boutures*

En analysant des  uvres pour la premi re partie (num riques ou non), nous nous sommes rendu compte que nous nous int ressions tout particuli rement   certains moments, parfois furtifs, dans des  uvres. Parmi d'autres exemples, un carr  qui monte et qui   un moment, se s pare en deux parties dans une  tude d'Oskar Fischinger. Nous pourrions,   la mani re du cin -concert<sup>767</sup>, poser notre propre bande-

<sup>767</sup> Des musiciens jouent (en direct ou sur un enregistrement) sur des films muets, abstraits ou non. Ils cr ent ainsi une bande-son qui n'existait pas dans l' uvre originale.

son sur ces s quences visuelles. Le principe n'est pas inint ressant, mais dans notre cas, nous souhaitons extraire des moments pr cis (visuels pour un film, sonores pour une musique), les reproduire par programmation informatique — les synth tiser — afin de pouvoir les manipuler, les cloner, et les confronter. Il s'agit en quelque sorte de boutures qui servent   faire rena tre des variations d'une m me souche : parfois en partant de sons (nous y greffons ensuite du visuel), parfois en partant de formes graphiques (nous y greffons du son). Dans le cas d'Oskar Fischinger, la s paration d'une forme en deux « morceaux » constitue une figure primaire. Nous l'observons au ralenti, et t chons d'en reproduire le comportement de mani re programm e. De cette mani re, il est ensuite possible de g n rer autant d'instances que voulu, de cr er des correspondances avec du son, et de rendre cela interactif, manipulable en temps r el. Nous exp rimentons ici une d marche de rem diation exp rimentale, assez proche de ce que propose le projet *ReCode*<sup>768</sup> par exemple.

### **Une compl tion. John Whitney sonifi **

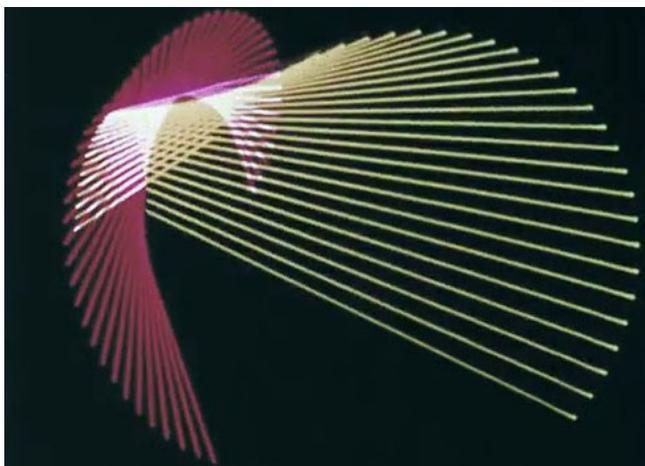


Fig. 203 : John Whitney, *Experiments in Motion Graphics* (1968)

<sup>768</sup> Le projet *ReCode* est un projet de rem diation d'œuvres pionni res de l'art par ordinateur. L'une des branches de ce projet s'appuie sur la revue *Computer Graphics and Art*, parue entre 1976 and 1978.   partir des images (et parfois m me des codes originaux) publi es dans la revue, des contributeurs re-programment ces pi ces (souvent   l' poque en FORTRAN) dans des langages informatiques plus r cents. Le code est mis   disposition et chacun peut proposer ses propres interpr tations ou variations. La vis e du projet est de faire revivre ces œuvres, de cr er une communaut  et d'offrir des ressources   destination de praticiens et d'enseignants (<http://recodeproject.com/>, consult  le 03/11/2022).

Lors de nos recherches sur John Whitney, nous avons constat  qu'un programmeur avait red velopp  une animation muette de John Whitney, et qu'il avait fourni son code<sup>769</sup>   la communaut . Nous avons retravaill  ce programme dans notre application, puis nous avons greff  des sons g n ratifs. Quand une ligne effectue un mouvement, nous changeons en m me temps des param tres sonores. Nous imaginons donc comment ces formes auraient pu  tre pens es de mani re audiographique.

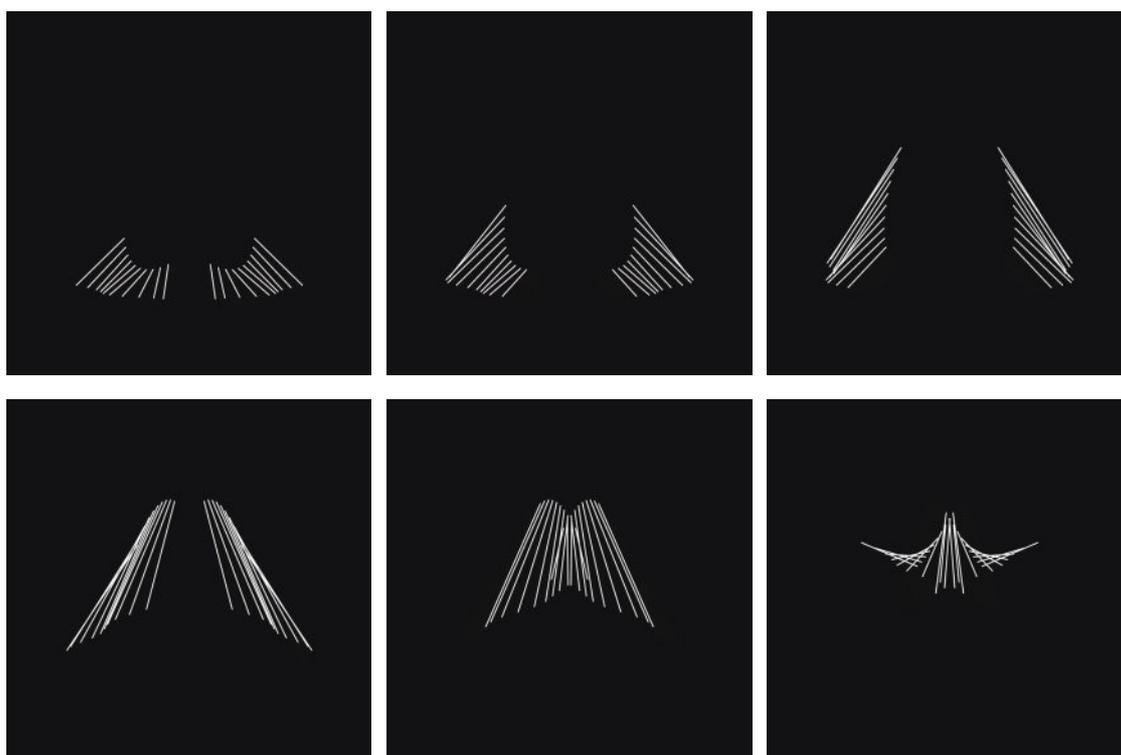


Fig. 204 :  tude audiographique d'apr s *Experiments in Motion Graphics* (1968), de John Whitney  
  audiographic lab

Dans cette  tude, un son saccad , de tonalit  descendante ou montante selon le mouvement est g n r . Un autre son est jou  en arri re-plan. Lorsque les deux parties de l'animation se croisent, ce son semble  tre associ    ce frottement. Nous pouvons  galement faire varier les param tres visuels et nous  loigner de la pi ce originale, voire laisser un utilisateur changer lui-m me les param tres.

769 Alexander Miller, « Recreating Vintage Computer Art with Processing », <https://www.youtube.com/watch?v=LaarVR1AOvs>, consult  le 30/01/23.

###   partir de *Study n  8* d'Oskar Fischinger : une  tude de scission

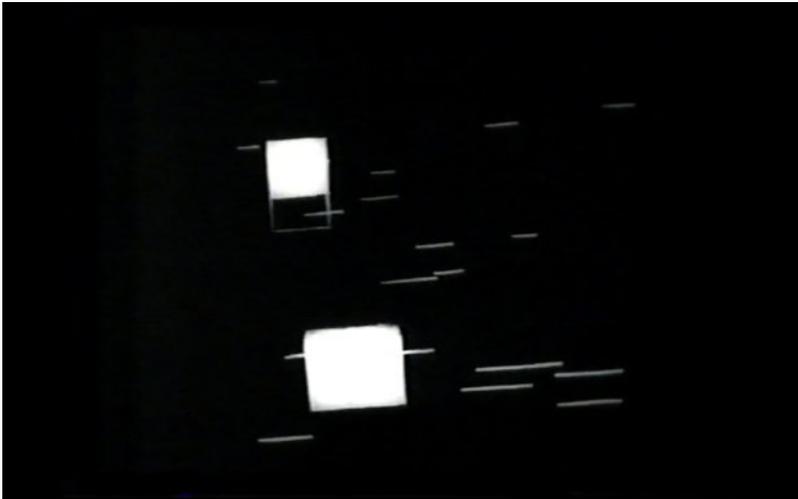


Fig. 205 : Oskar Fischinger, *Study n  8* (1932)  
  Center for Visual Music

Une image fixe rend difficilement compte de la ma trise du mouvement chez Fischinger. Dans ses  tudes en noir et blanc, c' st encore plus le cas. *Study n  8* (1932) utilise le po me symphonique *L'apprenti sorcier* (1897) de Paul Dukas.

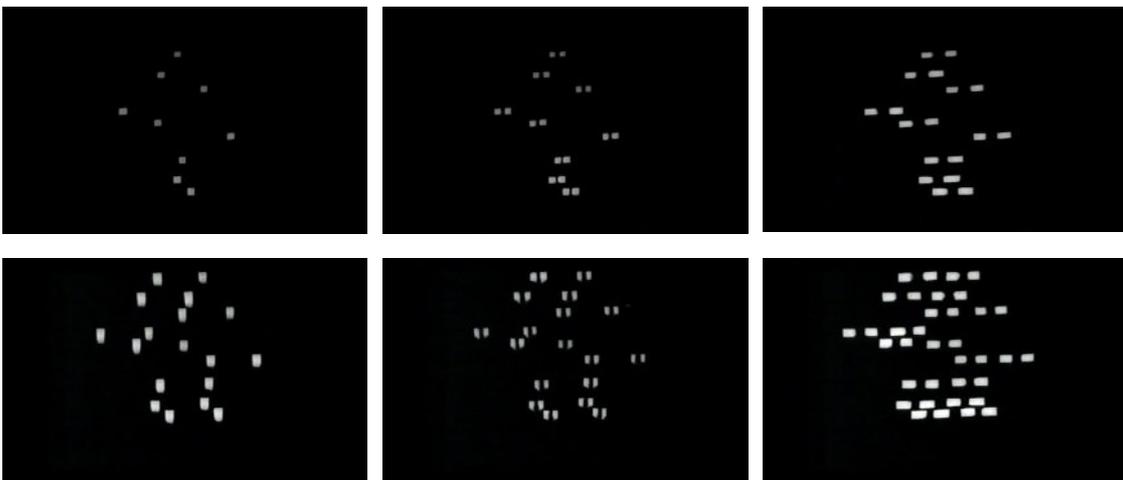


Fig. 206 : Oskar Fischinger, *Study n  8* (1932)  
  Center for Visual Music

Dans l'une des s quences, des formes apparaissent depuis le bas de l' cran et montent,   la mani re de bulles.   un moment donn , chaque forme se s pare, dans un mouvement de scission.

Comme nous l’avons expliqu  dans la premi re partie, Fischinger, alors m me qu’il avait invent  un syst me de synth se sonore, ne l’a pas utilis  dans ses films<sup>770</sup>. Pourtant, sa s rie d’ tudes (une dizaine de films, en noir et blanc) aurait constitu  un formidable terrain d’exploration. Prises isol ment, certaines s quences sont int ressantes en elles-m mes. Un moment notamment nous int resse beaucoup, nous l’avons  voqu  en d but de section : celui o  des formes quasi carr es montent vers le haut de l’ cran, et se s parent en deux morceaux juste avant d’atteindre leur objectif. Ces ensembles de scissions synchronis es nous paraissent d’autant plus int ressants que la transposition de cette figure d’animation en processus sonores — la scission d’un son en deux parties spatialis es — n’est pas quelque chose de courant.

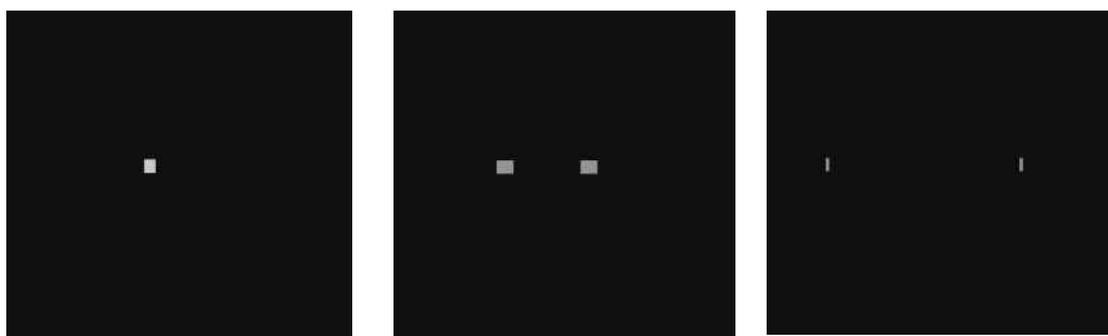


Fig. 207 : Reproduction programm e d’une scission, d’apr s un d tail de *Study n° 8* (1932) d’Oskar Fischinger  
   chaque phase de la scission, nous g n rons un son percussif diff rent  
   audiographic lab

Une fois l’isolation et la reprogrammation effectu es, on peut ensuite se demander si un son peut se « s parer ». En travaillant la spatialisation, cela peut s’av rer int ressant. En clonant ces formes, la question du groupe se pose  galement. Comme les formes montent, il y a aussi un mouvement g n ral, celui de l’ensemble.

<sup>770</sup> Sauf dans *Ornament Sound* (1932), mais qui est plut t un film-test, tr s diff rent de ses compositions anim es.

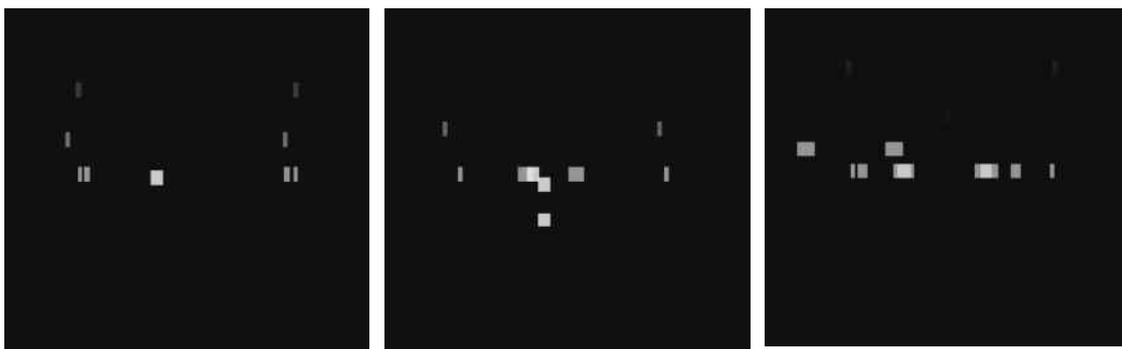


Fig. 208 : Une fois la scission recr ee sous forme d'objet informatique, nous pouvons la cloner. Dans notre  tude, une nouvelle scission est g n r e   chaque clic. Elle d bute   l'endroit o  est plac e le curseur de la souris. Lorsque de multiples scissions sonores sont pr sentes, la perception g n rale est celle d' clatements visuels et sonores incessants.  
  audiographic lab

Il s'agit donc de recrer,   partir de micro-s quences existantes, des situations  l mentaires, d'animation ou de son, pour ensuite les  puiser en explorant des variations : apparitions et disparitions, collisions, dur es, timbre, spatialisation...

Dans le cas de notre  tude de scission, cela peut aussi mener   d'autres pistes de travail, comme celle-ci (cf. ci-dessous), qui reprend le principe de mont e au clic, mais qui substitue   la scission la collision avec le haut de l' cran.



Fig. 209 : Dans cette  tude on clique pour g n rer une forme sonore. Le son associ    la forme d pend de la position de la souris. La forme se met   monter, et lorsqu'elle entre en collision avec le haut de l' cran, un son d cal  en fr quence se d clenche. Il est possible, en laissant appuy e le bouton de la souris, de « dessiner » des formes.  
  audiographic lab

### 4.3.3 Notions transversales. L'exemple d'une  tude de bruit

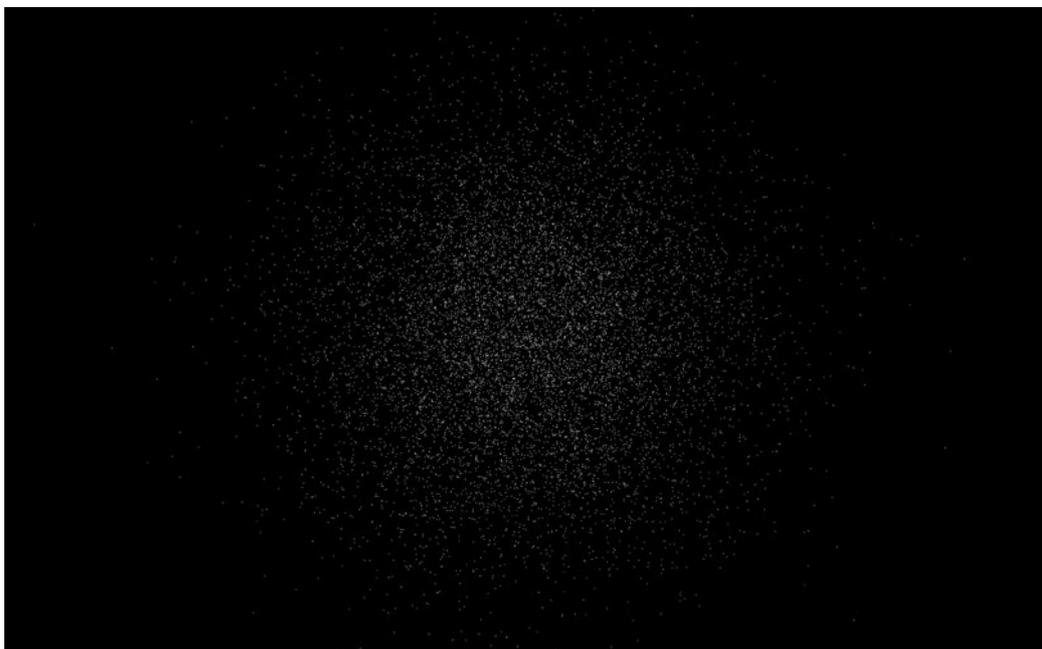


Fig. 210 :  tude de bruit graphique et sonore  
  audiographic lab

Une autre mani re de proc der pour la fabrication de nos  tudes consiste   explorer une notion commune au son et   l'animation, ou bien   transposer une figure existante de l'un des domaines dans l'autre. Nous prenons ici l'exemple de la notion de *bruit*, propre   la th orie de l'information, qui est utilis e   la fois dans le visuel et le sonore. Dans le sonore, elle d finit un son non musical, qui comprend un grand nombre de fr quences al atoires. Le « bruit blanc » est « un bruit dans lequel toutes les fr quences audibles sont susceptibles de se m langer<sup>771</sup> ». Lorsqu'on le produit synth tiquement, on peut en faire varier les param tres. Il existe ainsi des bruits color s : bruits roses, ou bruits marron, qui sont des variations particuli res du bruit blanc. Dans le visuel, le bruit traduit une r partition al atoire dans l'espace, et, par extension   l'animation, une agitation al atoire de points, comme pour la « neige » qui rempla ait l'image sur les  crans de t l vision lorsque l' mission du signal  tait d fectueuse.

---

771 Jean-Yves Bosseur, *Vocabulaire de la musique contemporaine*, op. cit., p. 28.

Dans notre  tude de bruit, la densit  de bruit (graphique et sonore) change al atoirement   chaque clic de souris. Lorsqu'elle est forte, le bruit sonore est plus important et la r partition de points blancs plus serr e, avec une agitation recentr e.

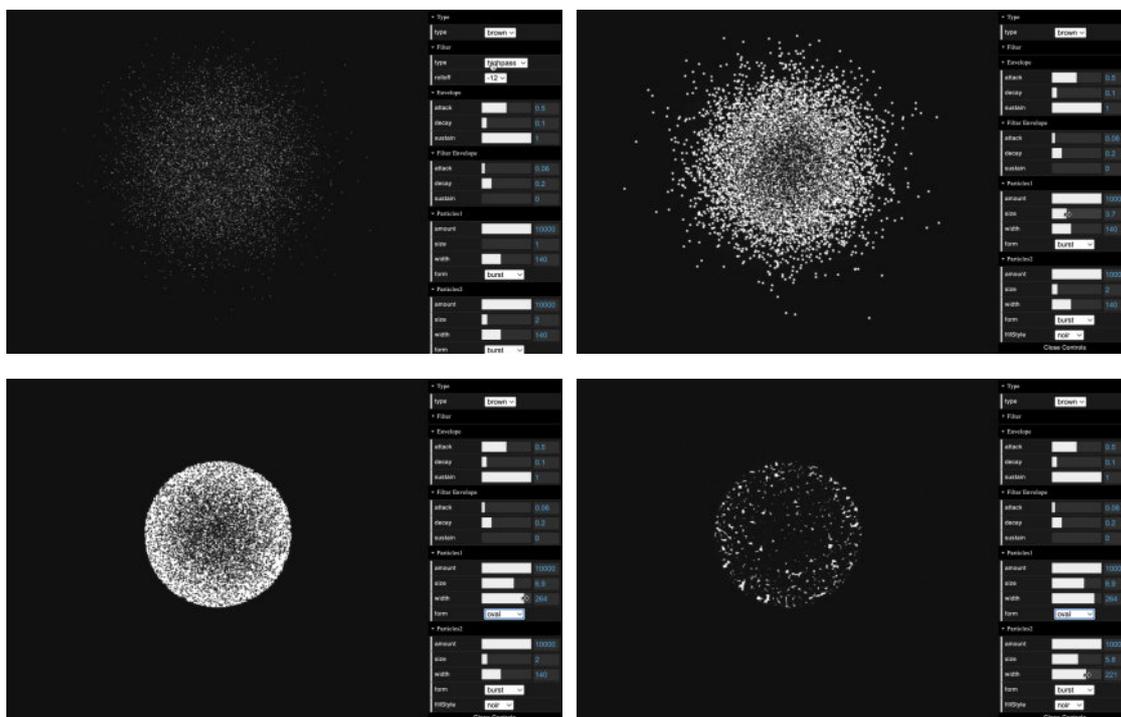


Fig. 211 : Diff rentes configurations de bruits graphiques et sonores  
  audiographic lab

Nous pouvons aussi, en superposant deux n buleuses de bruits, et en jouant sur le blanc sur noir, ou le noir sur blanc, cr er des formes plus complexes (cf. images ci-dessus). Se posent alors des questions de masquage visuel et de masquage sonore.

Chaque notion  voque bien entendu des  uvres artistiques, nous pourrions ici citer la composition *Analog #1 (Noise Study)* de James Tenney cr e en 1961 avec le logiciel MUSIC III aux Bell Labs, *Noisefields* (1974) de Steina et Woody Vasulka avec un travail sur le bruit vid o, le signal et le flicker.  galement le film structurel *Arnulf Rainer* (1960) de Peter Kubelka, dans lequel le cin aste n'utilise que des plans noirs ou blancs, du silence ou du bruit blanc. Plus r cemment, l'artiste et musicien Tristan Perich a r alis  un bo tier imitant celui du CD, sur lequel on vient brancher un casque (ce dispositif fait partie d'une s rie). La s rie que l'on y entend, appel e *Noise Patterns*

(2016), comprend 6 sections de variations de bruit blanc<sup>772</sup>. Toutes ces r f rences figureront en regard des  tudes correspondantes.

D'autres notions seront explor es de la m me mani re que le bruit, comme la trace et l' cho, le scintillement, l'interf rence, la permutation, le glissando...

#### 4.3.4 Approche perceptive : autour d'une illusion audiographique

M me avec une  tude basique comprenant simplement deux points anim s et deux sons, par exemple, se posent des questions de perception. Le cerveau cherchant   associer un son   une cause, il cherche la forme  mettrice de son la plus plausible. Dans un contexte d'abstraction visuelle, d gag  des principes du r el, on touche alors   des principes de Gestalt.

##### ***Le « Motion Bounce Effect » : l'effet de rebond***

De mani re tout aussi plus fondamentale, et dans un registre totalement abstrait, le *Motion Bounce Effect* met en sc ne deux disques qui se d placent vers le bas, en ligne droite diagonale, vers le coin oppos  chacun partant de deux coins gauche et droit d'une zone carr e. Si l'animation est muette, on per oit bien que les deux disques se croisent.



Fig. 212 : Reproduction de l'exp rience du Motion Bounce Effect  
  audiographic lab

Si l'on ajoute, au moment du croisement (au milieu de l' cran) un son bref, percussif, alors nous avons l'impression que les deux disques entrent en collision<sup>773</sup>. Ils semblent

772 Tristan Perich, *Noise Patterns* (2018), <https://tristanperich.bandcamp.com/album/noise-patterns>, consult  le 30/01/23.

773 La r f rence du premier article consacr    ce sujet est : Robert Sekuler, Allison B. Sekuler, Renee Lau, « Sound alters visual motion perception », *Nature* n  385, p. 308, janvier 1997.

rebondir l'un sur l'autre et d vier leur trajectoire, telles deux boules de billard qui s'entrechoquent<sup>774</sup>.

Dans l'une de nos  tudes, nous reproduisons par programmation ce comportement, et nous creusons exp rimentalement l'effet en dupliquant la paire de disques   diff rents endroits, avec ou non le m me centre de croisement.



Fig. 213 : Reproduction de l'exp rience du *Motion Bounce Effect*, avec plusieurs paires de disques qui se croisent  
  audiographic lab

Dans ce type d'exp rimentation perceptive, il est important que diff rentes personnes puissent faire l'exp rience de ces illusions. Nos exp rimentations  tant en ligne, il pourrait  tre possible de les faire tester par diff rentes personnes, en collaboration avec des laboratoires de recherche perceptive. Cela montrerait de mani re convaincante que le son peut influencer la perception du mouvement. Dans notre cas, il s'agit d'essayer de mieux comprendre comment notre perception visuelle et notre perception auditive fonctionnent ensemble. Comme nous associons un son   des ensembles de formes, une forme   un ensemble de son. L'apprentissage de la musique passe aussi par ces notions de perception, l'apprentissage du graphique aussi. Cette part de recherche pourrait  tre explor e plus en profondeur avec des structures sp cialis es.

Nous n'avons d taill  dans ce chapitre que quelques processus d'inspiration et de fabrication. Plusieurs  tudes sont d j  en ligne, d'autres sont pr tes    tre r alis es,   partir des diff rentes sources (manuels d' tudes, extraits de films ou de musiques...) que nous avons  voqu es au d but de cette partie.

<sup>774</sup> Une exp rimentation en ligne, de Michael Bach, dont nous inspirons, est accessible ici : <https://michaelbach.de/ot/aud-bounce/index.html>, consult  le 30/01/23.

## 4.4 Des interfaces : fa onnage, navigation et composition

Nous envisageons notre outil de mani re contrainte, pour rester   un niveau qui nous permette d'en ma triser les contours. Dans cette optique, la premi re  tape a consist    regrouper des ressources basiques,  l mentaires : extraits d'œuvres, manuels,  tudes sonores ou graphiques...   partir de ces ressources, d gager des pistes pour proposer des  tudes audiographiques interactives. Notre outil en ligne serait donc   la fois un moyen de naviguer dans des s ries d' tudes, de les exp rimer, de les annoter, en somme d'essayer des choses. Tester des correspondances, des oppositions, des contrastes, des variations : composer.

Cr ation d'une sorte de cours exp rimental, sur le mod le du Bauhaus pour l'image/son. L'id e est   la fois de publier r guli rement sous la forme d'un journal, tout en composant des formes de plus en plus complexes, et de d velopper autour des s ries de cours et ateliers d di s.

### 4.4.1 Navigation dans une page d' tudes



Fig. 214 : Menu d ployable pour naviguer dans les  tudes audiographiques  
  audiographic lab

Nous avons commenc    d velopper une taxinomie qui permettra de naviguer dans les  tudes. Celle-ci sera bas e sur des notions transversales ou des natures d' tude. Elle s'affinera au fur et   mesure du d veloppement.

Diff rentes pages d' tudes sont d j  accessibles. Une m me page peut regrouper plusieurs  tudes qui, dans l'ordre d'apparition dans la page (la consultation se fait pour le moment par d filement), font  tat d'une r flexion progressive autour d'une m me th matique ou de m mes contraintes. Un menu d ployable et repliable, positionn  en bas   droite, permet de passer de page en page.



Fig. 215 : Capture du d filement d'une page int grant plusieurs  tudes  
  audiographic lab

Sur une page li e   la mise en sc ne de plusieurs points et sons par exemple (ci-dessus la page « Points multiples »), je peux d'abord essayer dans le contexte d'une premi re  tude de tous les placer sur une horizontalit , puis d'essayer, dans une seconde, une trajectoire plus al atoire.

#### 4.4.2 Les trois modes de consultation d'une  tude

Chaque  tude poss de en elle-m me ses propres modes de visualisation et ses propres  l ments d'interface. Il existe trois modes principaux de pr sentation.

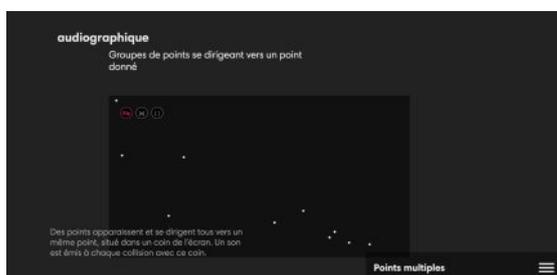


Fig. 216 : Le mode « page »  
  audiographic lab



Fig. 217 : Le mode « hybride »  
  audiographic lab

Le mode « page » pr sente l' tude sur une petite zone et autorise le d filement de la page pour passer d' tude en  tude afin de comprendre la progressivit . Le mode « hy-

bride »  tend l' tude sur toute la surface de la fen tre, mais conserve les textes explicatifs et les  l ments d'interface (cela peut permettre, entre autres fonctionnalit s, d'int grer des interactions entre le texte et l'animation). Le mode « plein  cran », enfin, fait dispara tre toute l'interface et permet de se focaliser sur l'exp rience, lin aire ou interactive. Il peut  tre utilis  dans le contexte mus al d'une installation, par exemple.

#### 4.4.3 Espaces de fa onnage

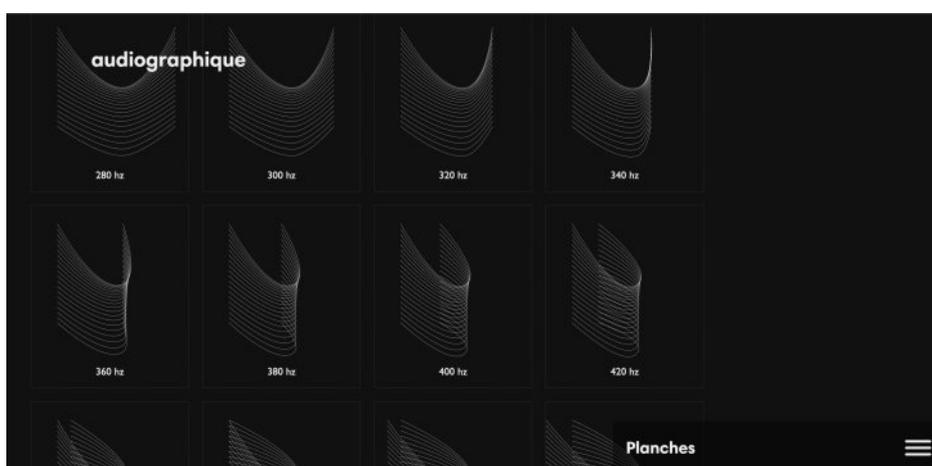


Fig. 218 : G n ration d'une « planche » qui correspond   la mise en correspondance du d veloppement d'une forme complexe (composite) avec une gamme de fr quences.  
  audiographic lab

Certaines formes graphiques ou sonores sont plus complexes que d'autres et n cessitent une sorte de fa onnage pr alable. Il s'agit principalement de tester et d'ajuster des param tres (dimensions, timbre, correspondances...) avant de les rendre disponibles dans la biblioth que, pour une mise en sc ne ult rieure dans une  tude publi e. Il s'agit   cette heure d'une interface qui nous est destin e, mais qui pourrait  tre   terme propos e   l'utilisateur.

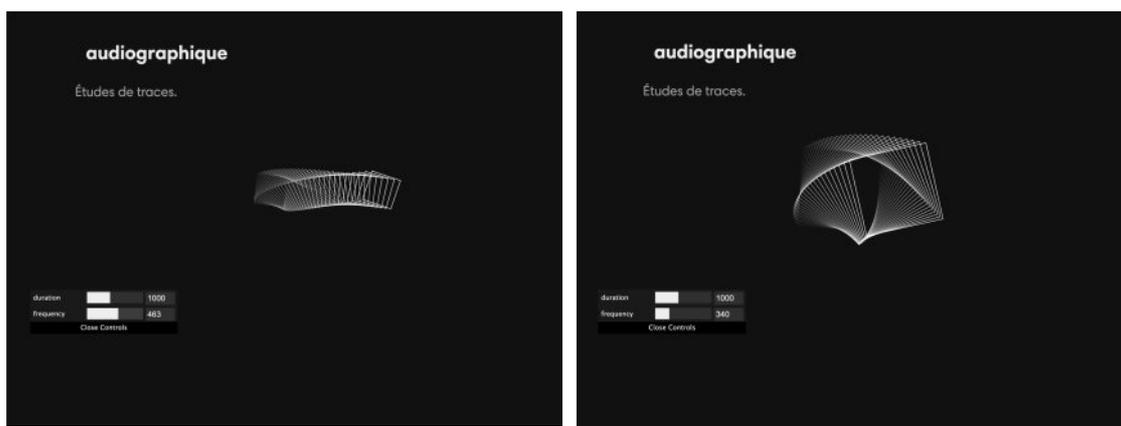


Fig. 219 : Fa onnage d'une entit  audiographique pour une  tude de trace  
  audiographic lab

Ces espaces vont  voluer, nous comptons exp rimer différents types de g n ration et de modelage.

#### 4.4.4 Une *timeline* de composition

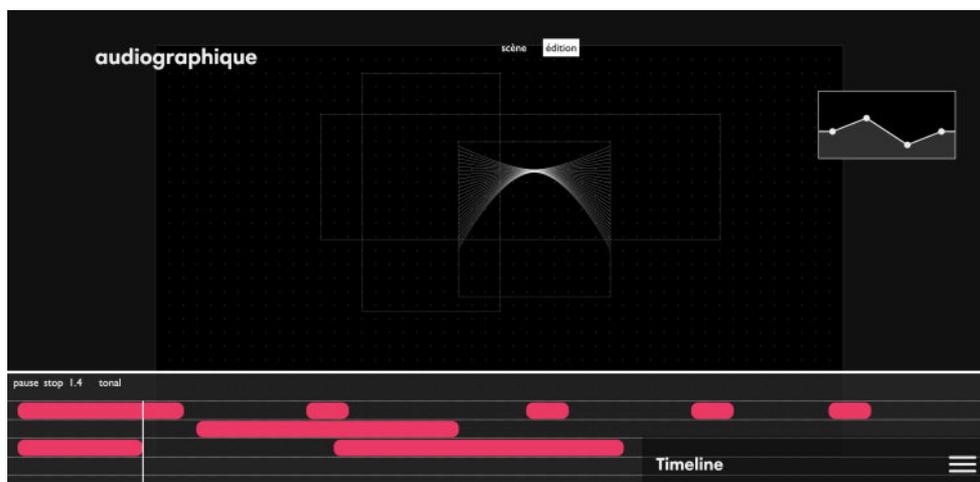


Fig. 220 : Capture de l'interface de composition, en mode « timeline »  
  audiographic lab

Au-del  des  tudes, nous souhaitons que ces exp rimentations nous m nent   composer des formes plus complexes, sortes de paysages audio-graphiques mall ables. Nous avons donc commenc    travailler sur une premi re interface de composition (nous en int grerons d'autres), qui utilise des principes assez standards de sc ne et de timeline, mais dans avec des objets audiographiques.

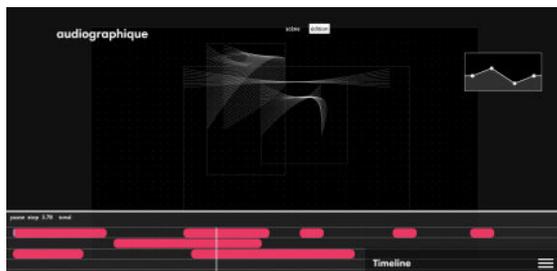


Fig. 221 : Mode «  dition »  
  audiographic lab

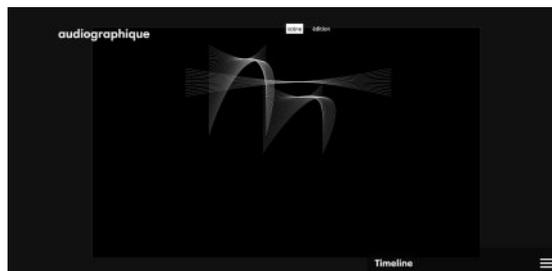


Fig. 221 : Mode « visualisation »  
  audiographic lab

Chaque piste correspond   une entit  audiographique, dont on peut configurer la position, les dimensions, le son. Il est possible  galement de d finir ou non un des contextes d'harmonie ou de tonalit  g n rale pour les sons.

#### 4.4.5 Exp rimentations sur les interfaces

Des interfaces comme la timeline sont assez r pandues. Toutefois, notre application  tant bas e sur les technologies web, nous envisageons, au-del  des  tudes, d'exp rimer en parall le sur les interfaces elles-m mes.

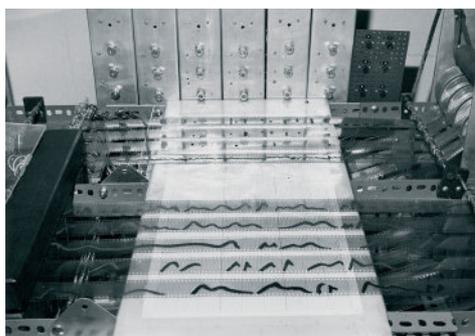


Fig. 223 : Le syst me *Oramics*, de Daphn  Oram (1957),   pointculture.be



Fig. 224 : Amit Pitaru, *Sonic Wire Sculptor* (2003)  
  YouTube d'Amit Pitaru

  titre d'exemple, des exp rimentations du dessin comme outil de g n ration font  cho   diff rents processus ou syst mes : parmi eux le dessin de « sons anim s » du cin aste canadien Norman McLaren, qui dessinait dans les ann es 1950 des motifs gra-

phiques sur la bande son de la pellicule<sup>775</sup> ou bien Daphn  Oram et son syst me *Oramics* (1957), qui lui permettait de produire des sons synth tiques en dessinant des formes sur une plaque de verre. Plus  volu , l'*UPIC* de Iannis Xenakis consistait   dessiner des formes plus ou moins complexes sur une tablette graphique. Une t te de lecture scannait ensuite de gauche   droite le dessin, convertissant les formes en sons, selon leur forme et leur positionnement. Le *Sonic Wire Sculptor* (2003), d'Amit Pitaru, fonctionne sur un principe similaire, mais gr ce au trac  par l'utilisateur d'un dessin filaire en 3 dimensions. En temps r el, ce dessin est lu comme une partition et g n re du son, en fonction des positions dans l'espace des trac s. Un trait plac  en hauteur sera par exemple aigu, et bas plus grave. Les ondulations et les dimensions de la zone du trac  jouent sur l'enveloppe sonore et la tonalit .

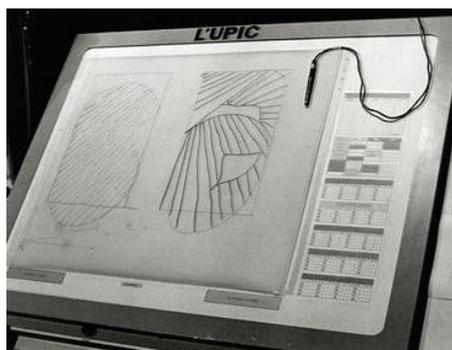


Fig. 225 : Le syst me *UPIC*  
  ResearchGate Sylviane Sapir / Centre Iannis Xenakis

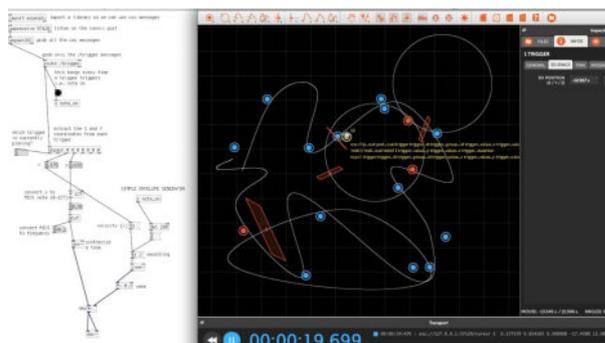


Fig. 226 : Le logiciel *Iannix*  
Ici la partition graphique interactive est branch e sur un patch sonore cr   dans *Pure Data*, environnement de programmation graphique,  iannix.org

La plupart des exp riences cit es ont d'ailleurs  t  rem diatis es<sup>776</sup>, voire repens es pour donner naissance   de nouveaux types d'applications. Par exemple, le logiciel *Iannix* s'inspire de l'*UPIC*. Son interface offre la possibilit  de dessiner ou de g n rer des formes graphiques, et d'y placer aux endroits voulus de petits marqueurs. Lorsqu'on lance le processus, le trac  sera consid r  comme une partition interactive. Chaque trac  est parcouru par une t te de lecture,   une vitesse param tr e par l'utili-

775 Ces principes sont expliqu s dans le documentaire de l'ONF, *  la pointe de la plume* (1951), [https://www.onf.ca/film/a\\_la\\_pointe\\_de\\_la\\_plume/](https://www.onf.ca/film/a_la_pointe_de_la_plume/), consult  le 30/01/23.

776 Des simulations num riques du son dessin  de McLaren, ainsi que d'*Oramics* de Daphn  Oram sont disponibles sous forme d'applications. L'application d'Amit Pitaru  galement, <http://www.sonicwiresculptor.com/>, consult  le 30/01/23.

sateur, et chaque rencontre avec un marqueur peut provoquer l'ex cution d' v nements. Nous utilisons le terme  v nement, car sans connexion avec d'autres logiciels, Iannix tourne « dans le vide ». Il ne g n re pas de son ni d'image directement, mais il programme des s ries d'ex cutions temporelles potentielles,   des conditions, gr ce   la cr ation d'une partition interactive non lin aire. Nous souhaitons  galement profiter des sp cificit s des interfaces web : le d filement, les cartes d pla ables, le paralaxe...

En parall le de ces interfaces graphiques, nous aimerions  galement mettre en place un ensemble de figures temporelles qui pourraient  tre directement appel es en langage naturel, sur le mod le des UST<sup>777</sup> (Unit s S miotiques Temporelles) d velopp es par le laboratoire de Musique et Informatique de Marseille (MIM) et inspir es par les travaux de Pierre Schaeffer<sup>778</sup>.

---

777 Sous la direction de Fran ois Delalande, des chercheurs ont rassembl  un certain nombre de « figures sonores » qui leur semblent r currentes ou remarquables dans des  uvres  lectroacoustiques. Elles sont divis es en deux cat gories principales. La premi re d finit les UST non-d limit es dans le temps : « en flottement », « en suspension », « lourdeur », « obsessionnel », « par vagues », « qui avance », « qui tourne », « qui veut d marrer », « sans direction par divergence d'information », « sans direction par exc s d'information », « stationnaire », « trajectoire inexorable ». La seconde cat gorie d finit les UST d limit es dans le temps : « chute », « contract - tendu », «  lan », «  tirement », « freinage », « suspension-interrogation », « sur l'erre ». Pens es au d part pour la musique, les UST se sont  galement d velopp es comme potentiel outil de caract risation d' uvres d'art appartenant   d'autres champs artistiques (Fran ois Delalande, « Les Unit s S miotiques Temporelles : probl matique et essai de d finition », in *Les Unit s S miotiques Temporelles,  l ments nouveaux d'analyse musicale*, Marseille, MIM, 1996, p. 18-19).

778 Pierre Schaeffer propose, dans son *Trait  des objets musicaux*, une s rie de sept crit res de valeur destin s   d crire un « objet sonore » : la masse, la dynamique (intensit  des diff rentes composantes), le timbre harmonique (couleur du son), le profil m lodique ( volution temporelle du spectre global), le profil de masse ( volution des composantes spectrales internes), le grain (irr gularit s de surface), et enfin l'allure, c'est- -dire l'analyse des vibratos (Pierre Couprie, « Les th ories de Pierre Schaeffer », <http://www.olats.org/pionniers/pp/schaeffer/theorieSchaeffer.php>, OLATS, 2000, consult  le 30/01/23, et Pierre Schaeffer, *Trait  des objets musicaux*, op. cit.).

## 4.5 Une version d di e au jeune public : *musigraphicode*

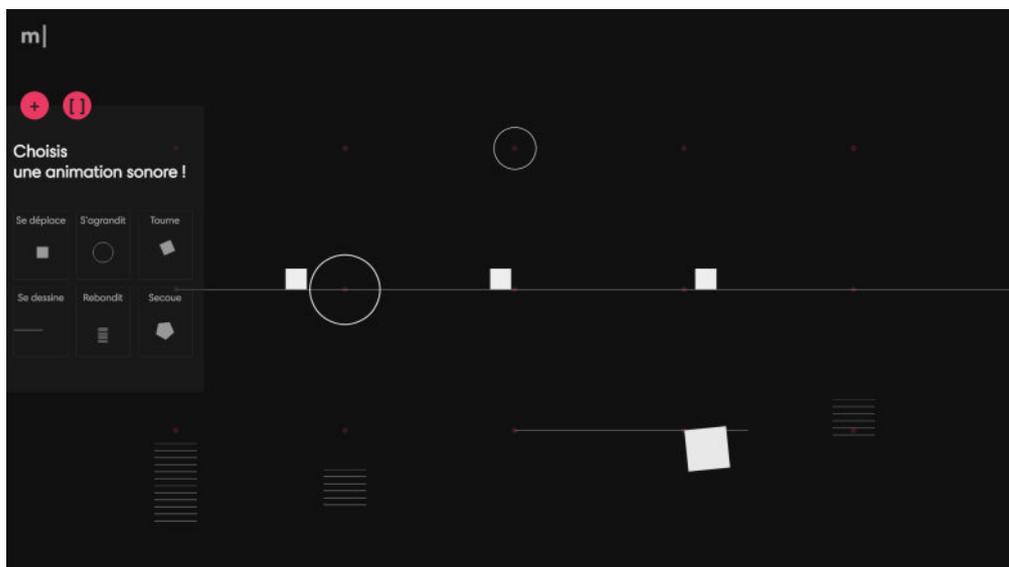


Fig. 227 : *musigraphicode*, capture de l'interface de cr ation et de composition  
  audiographic lab

### 4.5.1 Un atelier d'art audiographique et des projections

*musigraphicode* est une  dition particuli re de notre application, orient e pour un jeune public. Elle a  t  r alis e et pr sent e au Centre Pompidou en 2018,   l'occasion conjointe de la f te du code cr atif et de la Biennale N mo, d di e aux arts num riques contemporains. Diff rents artistes  taient pr sents, pour pr senter des applications originales, ou bien des versions adapt es de programmes ou logiciels qu'ils ont d velopp s pour des projets artistiques. Parmi eux, Antoine Schmitt, Christa Sommerer et Laurent Mignonneau, Anouck Boisrobert et Louis Rigaud, ou bien encore Christian Delecluse. Des performances audiovisuelles, comme *Tempest* de Frank Vigroux et d'Antoine Schmitt,  taient  galement programm es en parall le.

Pour tous les artistes invit s<sup>779</sup>, l'id e centrale  tait d'organiser un atelier autour de son application. Nous avons restreint, pour notre version pr sent e, les possibilit s de mani re   ce que des enfants puissent intuitivement comprendre le fonctionnement de l'interface. Nous en avons r sum  les possibilit s ainsi : « L'application web donne   l'enfant diff rentes possibilit s de g n rer des formes graphiques abstraites et des mo-

<sup>779</sup> Nous remercions Odile Fayet, cheffe de projet au service de la m diation Culturelle de la direction des publics du Centre Pompidou, qui nous a permis de pr senter notre travail.

tifs sonores puis de faire varier leurs param tres en direct. La vibration, l'ondulation, l' cho ou bien le bruit lui servent de guide pour cr er des tableaux anim s et sonores ». L'application est  galement accessible en ligne<sup>780</sup>.



Fig. 228 : *musigraphicode*, F te du code cr atif, Centre Pompidou, 2018

Sur la premi re image, au fond, l'espace de projection de films d'animation exp rimentale, ici *Symphonie Diagonale* (1921-24), de Viking Eggeling   Beno t Montign 

Nous disposons  galement d'un espace de projection   plusieurs entr es, dont l'une  tait accessible depuis l'espace de l'atelier. En collaboration avec le service de la m diation culturelle du Centre, nous avons assur  le commissariat d'une s lection d'œuvres issues des collections de films et vid os du Mus e national d'art moderne, entre musique visuelle et animation exp rimentale, parmi elles : *Symphonie Diagonale* (1921-24) de Viking Eggeling, *Filmcomposition* (1922) de Werner Graeff, *Dots* (1940), *Hoppity Pop* (1946) et *Synchromy* (1971) de Norman McLaren. Les films visibles  taient visibles depuis l'atelier.

780 <http://www.audiographiclab.com/musigraphicode/>, consult  le 30/01/23.

## 4.5.2 Fonctionnalit s de *musigraphicode*

*musigraphicode* permet   l'enfant de g n rer des formes graphiques abstraites, anim es et sonores. Une dizaine de postes  tait propos e dans l'espace de l'atelier. Chaque poste est un ordinateur sur lequel est lanc  un navigateur web en plein  cran. Deux haut-parleurs st reo sont associ s   chaque ordinateur. L'interface de l'application est compos e de la sc ne principale (un fond noir sur lequel apparaissent et s'animent les formes) et de deux zones d'interaction : la palette d'insertion d'animation, et la palette de modification des param tres par du code. Les deux palettes sont masquables une   une   tout moment, afin de laisser plus d'espace   l'arri re-plan.

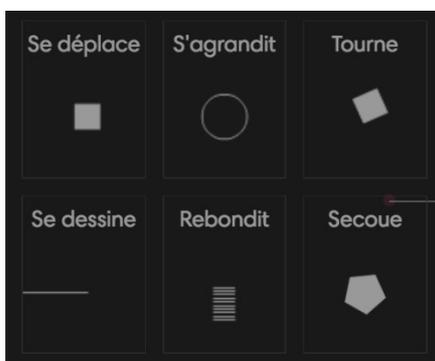


Fig. 229 : La palette d'insertion d'animations pr d finies  
  audiographic lab

Au d part la sc ne est vide. Seuls se trouvent neuf points rouges dispos s en grille ; ils symbolisent les emplacements disponibles pour les animations. La premi re palette,   gauche de l' cran, propose sous forme d'ic nes anim es six familles d'animations audiographiques : « se d place », « s'agrandit », « se dessine », « l vite », « tourne », « rebondit ». Chaque clic sur l'ic ne ajoute sur la sc ne l'animation correspondante,   un emplacement pseudo-al atoire — les entit s se positionnent en priorit  sur les espaces vides — sur la grille. Les nouvelles animations se calent automatiquement sur une mesure temporelle g n rale, pr programm e, et  voluent sur des dur es qui sont des multiples ou des fractions de cette mesure, comme en musique. Si on d passe neuf formes, alors les premi res apparues sont  cras es, de fa on   ce que l'on ne puisse

toujours avoir que neuf animations en m me temps<sup>781</sup>. Le r sultat d'ensemble est une composition audio-graphique tonale et harmonique.

### 4.5.3 Tonalit , rythmique, interchangeabilit 

Une fois que des animations sont pr sentes   l' cran et « entendables » sur les haut-parleurs, l'enfant peut, en temps r el, faire varier leurs diff rents param tres gr ce   une interface de *Live Coding*<sup>782</sup> d di e.

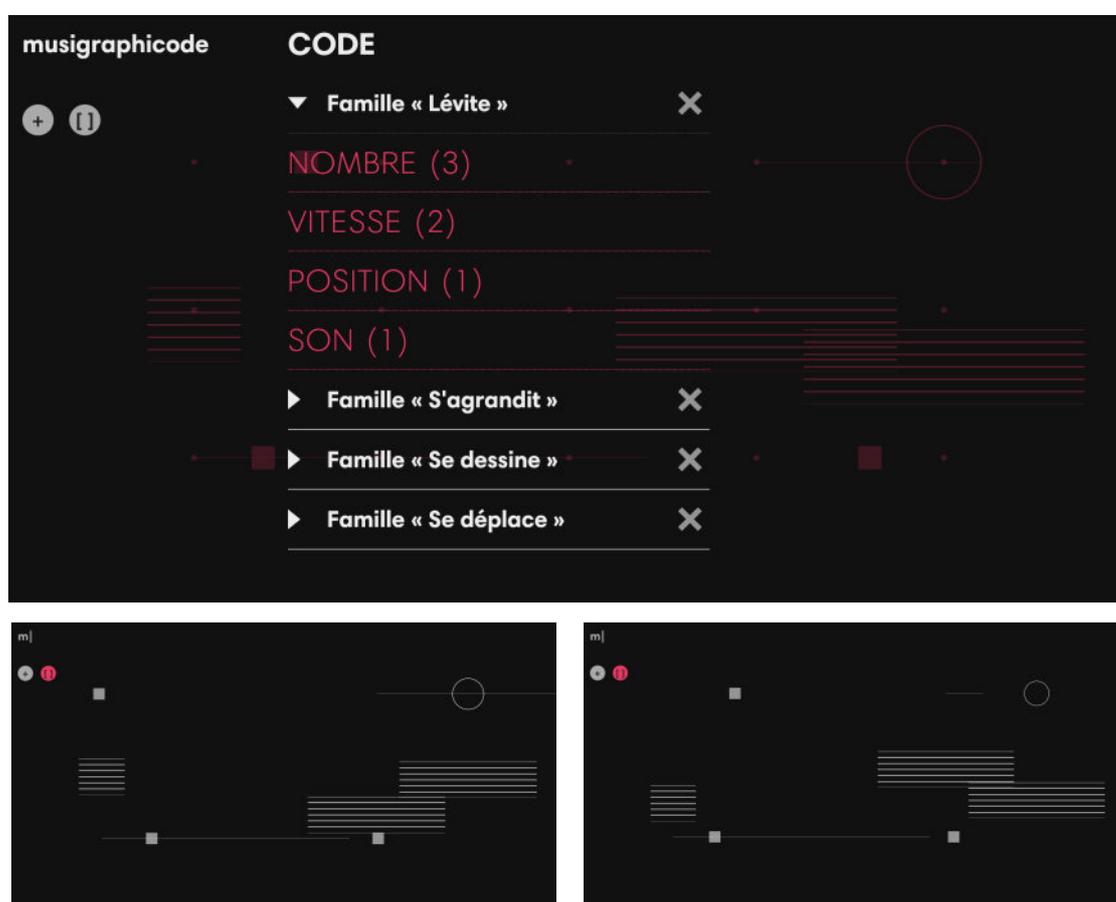


Fig. 230 : En haut la palette de *Live Coding*, en bas des exemples d'animation  
  audiographic lab

781 Beaucoup d'enfants ont tendance, au d but de l'exp rience interactive,   cliquer sans arr t. Nous utilisons cela pour reconstituer sans cesse une nouvelle composition.

782 Un artiste-programmeur modifie en direct le programme central (qui g n re des formes) gr ce   un  diteur de code visible par les spectateurs. Les r percussions graphiques ou sonores sont imm diates, cela permet au public de mieux appr hender les manipulations effectu es.

Chaque famille, si elle est repr sent e   l' cran, dispose d'une zone de param trage d di e. Toutes ces zones sont d ployables et repliables, et l'on peut faire d filer l'ensemble pour passer de l'une   l'autre. Pour chaque famille, il est possible de changer : le nombre d'instances pr sentes, leur vitesse d'animation, leur position et le son associ . L'enfant peut ainsi essayer plusieurs sons sur une m me forme graphique. Des sons plus ou moins m talliques, des nappes plus ou moins douces, des sons percussifs.

#### 4.5.4 Des correspondances « primitives »

Nous utilisons dans cette exp rience des correspondances intuitives entre graphique et sonore, assez proches de celles point es par le psychologue Albert Wellek dans les ann es 1920<sup>783</sup>. Plus la forme est positionn e sur le haut de l' cran, plus la tonalit  du son qui lui est associ  est aigu . Plus bas, elle sera plus grave. Une correspondance spatiale existe aussi entre position horizontale et spatialisation sonore. Si une entit  se d place de gauche   droite, on entendra le son aller de gauche   droite dans les enceintes, comme dans l'une de nos  tudes audiographiques appartenant   la s rie « Un point, un son<sup>784</sup> ». En changeant en direct la position des entit s, l'enfant fait donc aussi varier la tonalit . Un positionnement « au hasard » est aussi propos . Il est aussi possible de supprimer directement toute une famille. L'id e de l' v nement est d'initier les enfants au code de mani re ludique. Dans notre interface, si le code est mal  crit, il se colore en gris et ne fonctionnera pas. Si la syntaxe est bonne, il se met en rouge et cela a un impact direct ; on voit soudain les formes s'agiter, on entend des changements de son. Les diff rents sons musicaux ont  t  choisis pour s'accorder entre eux,   partir d'une tonalit , simple, en Do majeur. Certains sons sont plus percussifs, ils viennent cr er des rythmes.

Les possibilit s sont ici volontairement restreintes, mais le logiciel est capable d'accueillir une base illimit e de motifs sonores et graphiques. En effet ceux-ci ne sont pas constitu s de fichiers m dia lourds, mais de simples fichiers textes l gers qui ne contiennent que des suites de param tres sonores ou graphiques.

---

783  tudiant les connexions entre le visuel et le sonore, Wellek avait d gag  six correspondances fondamentales — des « synesth sies primitives ». Pour lui, on associe par exemple plus fr quemment les formes minces   des sons aigus, les formes  paisses   des formes graves (Gehard Daurer, « Audiovisual perception », in *See this sound, op. cit.*, p. 333).

784 Voir le chapitre 4.3, consacr    la description d'une s lection de nos  tudes.

Dans cet environnement rudimentaire, il est déjà possible, en attribuant des sons et des vitesses à une famille d'éléments, et d'autres paramètres à une autre, de voir si les formes se distinguent, ou si tout devient confus. Il s'agit pour nous de questions essentielles.

## Conclusion

### L'espèce « cinéma sonore » à l'ère du numérique

Pour tenter de comprendre les relations qui se nouent entre les images et les sons à l'ère du numérique, nous avons pris le parti de définir trois grands axes de réflexion : la musique visuelle programmée, les nouvelles dramaturgies interactives qui s'appuient sur des bases de données, et un ensemble de travaux artistiques mettant en jeu le visage et la voix. Ceux-ci prennent leur source à la période de stabilisation de la synchronisation audio-visuelle dans le cinéma sonore des années 1920. Le premier axe a trait à l'idée de la *musique visuelle*, c'est-à-dire un champ artistique hybride au sein duquel des correspondances s'exercent entre le musical et l'abstraction visuelle en mouvement. Le deuxième axe est celui de l'étude du « réel saisi par les machines<sup>785</sup> » : les images en mouvement et les sons peuvent être capturés, même séparément, pour être inscrits et réunis sur un même support, la pellicule cinéma. Le troisième axe, enfin, traite de la question des visages, des voix et du dialogue, à partir de l'émergence du cinéma parlant. Les chapitres « pré-numériques », qui introduisent chaque partie, malgré leur brièveté, cherchent quant à eux à démontrer l'utilité de croiser des histoires souvent cloisonnées, celles des sons et des images, dans leurs aspects esthétiques et techniques.

Il n'a toutefois pas été question de dessiner, à partir de ces trois champs, une *grande* histoire, exhaustive, mais plutôt de porter ceux-ci comme lignes directrices au moment de s'immerger dans le milieu complexe du champ numérique. Comme nous l'avons en effet évoqué à plusieurs reprises, les relations entre image et son jusqu'au milieu du XX<sup>e</sup> siècle — voire même jusqu'à l'art vidéo<sup>786</sup> — ont fait l'objet d'études relativement étendues. Il en est de même pour le son à l'image dans le cinéma dit « traditionnel ». En revanche, les analyses audio-visuelles, à partir — et autour — de l'émer-

---

785 Nous reprenons l'expression employée lors d'un entretien mené par Jean-Louis Boissier et Jean-Marie Dallet avec Piotr Kowalski (Jean-Louis Boissier, Jean-Marie Dallet, entretien avec Piotr Kowalski, *op. cit.*).

786 Des ouvrages comme celui de Holly Rogers (Holly Rogers, *Sounding the Gallery: Video and the Rise of Art-Music*, Oxford, Oxford University Press, 2013) ou celui de Paul Hegarty (Paul Hegarty, *Rumour and Radiation. Sound in Video Art*, Londres, Bloomsbury, 2015) sont par exemple spécifiquement consacrés à la question du son dans l'art vidéo.

gence des nouveaux média, sont, elles, beaucoup plus rares. Même si « l'espèce cinéma<sup>787</sup> » a subi une mutation numérique, elle est cependant restée foncièrement fidèle à ses origines. Le numérique a permis d'y intégrer davantage d'effets de synthèse, d'en faciliter la spatialisation sonore, mais elle n'est plus la forme audio-visuelle dominante (et cela même probablement depuis la télévision). Elle a vu se développer dans son giron une multitude de modèles non linéaires. Nous pouvons alors considérer — dans le prolongement de ce que suggère Jean-Marie Dallet pour le couple image interactive/informatique<sup>788</sup> — que le cinéma traditionnel (linéaire, à 24 images par seconde), point de départ de notre étude, est devenu un cas particulier parmi l'ensemble des formes audiovisuelles qui circulent dans la sphère contemporaine. Aussi il nous a paru potentiellement fructueux d'adopter, au-delà des champs disciplinaires de la musique et du cinéma, une approche pluridisciplinaire.

Dans les trois premières parties de notre travail, nous avons donc tâché de réunir des œuvres, à caractère numérique, dans lesquelles le son et l'image dialoguent dans un rapport d'équité. Le prisme audio-visuel nous a semblé être une approche analytique pertinente. Il permet à la fois une appréhension du milieu numérique, mais tout aussi bien une mise en perspective historique et transversale.

## Des correspondances à l'apprentissage automatique

La première partie, consacrée aux musiques visuelles, retrace la manière dont le numérique a permis de réaliser le rêve des artistes qui tentaient d'établir des correspondances entre musique et peinture : réunir la manipulation de formes graphiques abstraites et de sons grâce à un même dispositif, qui plus est en temps réel. Il en résulte des formes audio-graphiques génératives et interactives, muséales ou performatives, à la frontière du musical et de l'animation graphique. Il ne s'agit toutefois plus d'analogies chimériques entre couleur et son, qui existeraient dans la nature, mais de mappings de paramètres établis librement par les artistes. Dans ce domaine, les créateurs œuvrant avec le numérique tombent fréquemment dans l'écueil d'une conversion di-

---

787 « Une fois posée ce qu'est la forme cinéma, on observe que les dispositifs cinématographiques interactifs ont joué avec sa définition, créant de nouvelles branches de l'espèce cinéma. Des modèles historiquement abandonnés vont être réutilisés : stéréoscopie, panorama, mobilité du projecteur, variabilité de la position du spectateur, etc. » (Jean-Marie Dallet, « Un cinéma de la base de données », in Caroline Renard (dir.), *Images numériques, op. cit.*, p. 227).

788 *Ibid.*

recte d'image en son, ou bien de son en image<sup>789</sup>, considérant que ce processus constitue une œuvre en soi. Or une telle traduction systématique, isolée, est vide de sens. Il nous semble que tout rapport entre image et son doit être pensé, creusé, exploré. C'est ce même fondement que l'on retrouve dans les travaux des artistes étudiés dans la première partie. Avec John Whitney et son logiciel de composition, il s'agit, à partir d'outils techniques communs au visuel et au sonore, d'inventer un langage de composition audiographique. Ryoji Ikeda, lui, cherche à mettre au point un langage radical d'hyperméchronisation audiovisuelle, qui relève d'une démarche tout à fait contemporaine. Ces possibilités de traductions directes et instantanées ont tout de même fait naître des formes de visualisation et de sonification, qui ont comme base commune des réalités représentées par des données numériques.

La deuxième partie, dédiée aux médias numériques faisant appel à la saisie du réel, montre que l'interactivité associée à l'instantanéité de l'accès à des bases de données a produit un déplacement du rôle du monteur. Celui qui, selon Vilém Flusser, joue avec la linéarité : répétition, inversion du déroulement, saut dans le temps, accélération, collage et boucle temporelle<sup>790</sup>. Désormais, toutes ces possibilités sont transférées à l'utilisateur qui interagit avec l'œuvre, dans des environnements de plus en plus habités par les bases de données et les archives. Leur prolifération entraîne un excès de mémoire, propre à notre époque, c'est ce que dénonce N. Katherine Hayles<sup>791</sup>.

Enfin, dans la troisième partie consacrée à la voix et au visage, nous constatons que le numérique s'insinue à de multiples endroits. Il a ouvert la porte à des portraits parlants qui se manifestent tels des fantômes, captures de l'existant générées artificiellement. Comme le fait remarquer Anne Beyaert-Geslin, face à la prolifération des *selfies*

789 C'est-à-dire utilisant les possibilités gyrovagues du numérique selon la formule de Bachimont. Nous rappelons que le philosophe emploie ce terme pour définir un caractère d'indifférence propre au numérique (Bruno Bachimont, « L'archive et la massification des données : une nouvelle raison numérique », *Gazette des archives* n° 245, 2017, [https://www.persee.fr/issue/gazar\\_0016-5522\\_2017\\_num\\_245\\_1\\_5510](https://www.persee.fr/issue/gazar_0016-5522_2017_num_245_1?sectionId=gazar_0016-5522_2017_num_245_1_5510), consulté le 30/01/23).

790 « Il [Le monteur] peut répéter les événements, en inverser le déroulement, sauter des phases comme le cheval saute des cases aux échecs, aller du futur au passé, accélérer le cours du temps ou le freiner, coller ensemble le début et la fin du temps linéaire et faire ainsi une histoire cyclique ; bref il peut jouer avec la linéarité » (Vilém Flusser, « De la production et de la consommation des films », in Vilém Flusser, *La civilisation des médias*, Belval, Circé, 2006, p. 79).

791 « L'hypermnésie ou l'excès de mémoire est propre à la culture du XXI<sup>e</sup> siècle. À mesure qu'augmentent les capacités de stockage de la mémoire, il faut de plus en plus de bases de données relationnelles pour organiser, traiter, rechercher et récupérer les données » (N. Katherine Hayles, « Non-conscient cognitif et algorithmes de *trading* automatisé », in Claire Larssonneur, Arnaud Regnaud, Pierre Cassou-Noguès, Sara Touiza (dir.), *Le Sujet Digital*, Dijon, Les presses du réel, 2015, p. 20).

et des images d'exposition de soi sur les réseaux sociaux, un tirage argentique de visage devient un artefact précieux, de par sa rareté<sup>792</sup>. Nos visages et nos voix sont devenus de simples codes qui servent à déverrouiller nos smartphones. Sujets à la reconnaissance automatique, et par là même aux biais identitaires (Hito Steyerl et Trevor Paglen), ils servent aussi à anthropomorphiser des *chatbots* ou des formes textuelles numériques comme les SCAMs (*The Rumour of the World*, de Joana Hadjithomas et Khalil Joreige).

## Figures émergentes de l'image-son

Même si nos trois parties couvrent trois champs d'exploration relativement distincts, des figures d'image-son transversales émergent. Parmi elles, certaines pratiques audiovisuelles préexistent au numérique et se renouvellent en étendant leur champ d'action : le réemploi ou la performance audio-visuelle. D'autres sont propres à l'ère numérique : les logiques de capture de données, de détection et de surveillance.

Le réemploi ou *remix* est un exemple de pratique préexistante, que l'on a pu voir revisitée : par le détournement numérique et la spatialisation chez Candice Breitz, ou bien encore par la semi-automatisation permise par la base de données et la requête, comme chez les SLIDERS\_lab. Encore plus en concordance avec notre époque, Cory Arcangel, Natalie Bookchin ou Christian Marclay réinventent ces pratiques par l'utilisation des *Vlogs* ou des *Snapshots*, ces tronçons d'images-sons issus des réseaux sociaux. Tout comme le réemploi, les formes scéniques — depuis les représentations musicales à l'écran du précinéma — se sont aussi hybridées : en témoignent les formes de *Live Cinema* et de performance audiovisuelle. Nous pouvons par exemple citer les concerts audiographiques hérités de la musique visuelle (Ryoji Ikeda, Golan Levin, Antoine Schmitt...), les dispositifs performatifs basés sur la navigation en temps réel au sein de bases de données (Granular Synthesis, Jérôme Joy, SLIDERS\_lab...), ou enfin des formes audiovisuelles de spectacle vivant (Alexander Schubert).

D'autres pratiques éclosent encore avec la mise en réseau : non plus des logiques de saisie audiovisuelle directe du réel, mais des processus de capture d'information, de données brutes. Même les images et les sons enregistrés (*Field Recording*) recèlent des

---

792 Anne Beyaert-Geslin, *Sémiotique du portrait. De Dibutade au selfie*, Paris, De Boeck Supérieur, coll. « Culture & Communication », 2017, p. 96.

métadonnées qui permettent à l'artiste d'inventer de nouvelles navigations (le GPS chez Masaki Fujihata). Décrivant aussi bien la vie de nos sols (les tremblements de terre dans chez Semiconductor), des flux éoliens (Peter Beyls), que nos battements de cœur (Jacopo Baboni Schilingi), ces nouvelles formes audiovisuelles prennent le pouls du monde. Elles se manifestent dans une sorte de *monitoring* d'images et de sons qui nous sensibilise à la fragilité de la nature et cherchent à nous faire éprouver ce qui nous est invisible ou inaudible. Dans le même temps, il s'agit pour les artistes de faire prendre conscience au grand public de l'impact environnemental de ces technologies soi-disant immatérielles. L'extraction de matériaux rares dans les mines (Moritz Fehr), et le développement croissant d'infrastructures de stockage et de circulation permettant aux média numériques d'exister (Trevor Paglen) ne sont pas sans conséquences. L'autre pendant inquiétant est celui du monde de la surveillance, de la détection et de l'analyse automatisée ; non plus pour surveiller la planète et garantir sa santé, mais pour surveiller les personnes, profiler, et développer des stratégies marketing.

Avec le numérique, les images et les sons sont capturés, indexés dans des bases de données. Leur accessibilité est immédiate. Ils ouvrent alors la porte aux technologies de surveillance, de détection, de localisation, de classement, de recommandation, propres à notre sphère numérique contemporaine. Images et sons se voient soumis aux mêmes détournements : ils deviennent capitalisables (la chasse à la voix), et diagrammatiques (ils deviennent des mots de passe). Si pour Christian Marclay l'analyse informatique du son permet l'opération d'un montage poétique par similarité sonore, pour les industriels cette même analyse est utilisée pour gagner en efficacité et en automatisation dans la surveillance (comme Hito Steyerl l'évoquait dans *The City of Broken Windows*, autour d'une société qui pratique l'analyse sonore de bris de verre).

Dans ce contexte, la confiance croissante dont bénéficie le sonore<sup>793</sup> tend à s'étioler avec l'avènement des assistants vocaux et des voix artificielles. À raison, car les dispositifs qui les font exister masquent en effet des écosystèmes de grande ampleur qui favorisent le travail précaire, la pollution numérique, la consommation excessive et anarchique d'images et de sons.

---

793 Nous pouvons citer l'exemple du procès des attentats du 13 novembre 2015, accessible en direct, mais seulement en audio, en ligne, afin d'éviter une « dramatisation » par le montage visuel.

## Ouverture de notre recherche-cr ation

Notre recherche-cr ation s'est majoritairement construite autour de la premi re partie, dans une exploration qui nous semble fondamentale : celle de la circonscription d'un art audio-graphique, mettant sur un pied d' galit  image et son. D s lors qu'une forme graphique anim e et un son sont associ s, il s'agit de penser leurs relations. Nous avons donc t ch  de croiser les basiques de la synth se sonore (fr quence, timbre, balance, rythme...) et ceux de l'animation graphique (le d placement de formes primaires comme le point et la ligne, leurs transformations, leurs apparitions et disparitions...). Cela s'est concr t s  dans la r alisation d' tudes basiques (variations de fr quence, synth se additive,  cho, contrastes, acc l ration, inversion...), amen es    voluer vers des compositions plus avanc es (paysages graphiques et sonores, contrepoints graphiques, mouvements sonores...). Toutes les  tudes et compositions *audiographiques* ont  t , dans cette premi re phase de travail, bas es sur des motifs graphiques en noir et blanc, afin que l'accent soit mis sur le mouvement des formes. Certaines figures transversales ont  merg  de ce travail, parmi elles : le bruit, la variation, la trace et l' cho, ou bien encore le glissement. Ces  tudes interactives d j  publi es soul vent  galement diff rents questionnements relatifs   la perception multimodale et aux illusions perceptives.

  pr sent, tout en souhaitant continuer ces s ries d' tudes, nous aimerions  toffer cet ensemble par des exp rimentations audio-visuelles guid es par les travaux d'artistes qui nourrissent les deuxi me et troisi me parties. Tout en conservant notre qu te d' quilibre entre l'image et le son, cette recherche pourrait ainsi se poursuivre par la r alisation d' tudes cette fois-ci *audio-vid ographiques* visant   explorer   la fois le c t  structurel, mais aussi l'aspect de la mat rialit  num rique, par le fragment audiovisuel, et par les visages-voix.   ce titre, les figures mentionn es pr c demment, issues des trois parties, pourraient constituer un prolongement naturel de notre recherche. Il en est de m me pour des notions telles que les questions de compression, de d finition, d'obsolescence, de stockage et de spatialisation des syst mes de diffusion.

## Face aux interfaces invisibles. Quelles stratégies artistiques ?

Nous nous sommes cantonnés à des formes écraniques et sonores, héritées du cinéma. Par conséquent, nous n'avons fait qu'effleurer les problématiques liées au jeu vidéo, à la réalité augmentée, à la réalité virtuelle, ou bien encore aux *Metavers*. Les relations entre image et son s'opèrent pourtant de manière singulière dans ces milieux en constante évolution. Les domaines liés à la *Blockchain*, notamment les NFT (Non-Fungible token), sont aussi de nouveaux terrains d'exploration pour les artistes. Ils sont en revanche sujets, pour le grand public, à une grande incompréhension, de par leur opacité. Il y a là l'idée récurrente de « boîte noire », c'est-à-dire un système fermé que l'on utilise sans comprendre son fonctionnement interne<sup>794</sup>. À ce titre, l'artiste pionnière du « net art » Olia Lialina rappelle une tendance actuelle chez les industriels du numérique : bannir progressivement l'utilisation du terme « utilisateur » au profit de termes plus vagues, comme celui d'« expérience ». Pour elle, il s'agit d'un danger au sens où, le concept d'*utilisateur* se rapportant précisément à l'utilisation consciente d'un dispositif informatique, le gommer revient à faire oublier la présence des systèmes programmés qui nous entourent<sup>795</sup>. Or, toutes les technologies actuelles concourent à faire oublier la machine, au risque de ne plus comprendre comment elles sont programmées. C'est notamment le combat de Douglas Rushkoff qui affirme la nécessité d'apprendre à programmer, pour ne pas risquer d'être nous-mêmes programmés<sup>796</sup>. Il encourage ainsi à ne pas enseigner l'utilisation des programmes — des logiciels clés en main —, mais bel et bien la programmation<sup>797</sup>, afin de mieux comprendre le fonctionnement de l'ordinateur.

Pour se soustraire à ces problématiques, les artistes imaginent différents stratagèmes pour lutter contre les diktats de l'appareil industriel qui impose en permanence, sous

---

794 Frank Pascuale, spécialiste de la gouvernance algorithmique, définit la boîte noire ainsi : « Nous pouvons observer ce qui y entre et ce qui sort, mais sans être capables de dire comment l'un devient l'autre » (Frank Pascuale, *Black Box Society. Les algorithmes secrets qui contrôlent l'économie et l'information*, Limoges, FYP Éditions, coll. « Présence/Essai », 2015, p. 12-13).

795 « [...] lorsqu'on l'interprète dans un contexte plus large, le désaveu du mot "utilisateur" devient dangereux. L'existence de l'utilisateur est la dernière trace qui nous rappelle la présence, qu'il soit visible ou non, de l'ordinateur, d'un système programmé que l'on utilise » (Olia Lialina, « L'utilisateur Turing-complet », trad. Jean-François Caro, 2013, <http://reader.lgru.net/texts/lutilisateur-turing-complet/>, consulté le 30/01/23).

796 « La programmation est le nerf de la guerre d'une société numérique. Si nous n'apprenons pas à programmer, nous risquons de l'être nous-mêmes » (Douglas Rushkoff, *Les 10 commandements de l'ère numérique*, FYP, coll. « Innovation », 2012, p. 159).

797 *Ibid.* p. 155.

prétexte de nouveauté, ses machines et ses formats. Comme l'explique la théoricienne Nicole Brenez, les créateurs pratiquent une sorte de désobéissance en inventant leurs propres outils, en détournant l'existant (le *hacking*) ou en réutilisant d'anciens systèmes<sup>798</sup>. Ils emploient la *ruse*, comme le préconisait Flusser<sup>799</sup>, pour combattre la prolifération de ce qu'il appelle à l'époque les « images sonorisées » techniques<sup>800</sup>. De manière plus générale, la situation pourrait être résumée par cette phrase d'Olia Lialina : « Nous aurions tort de penser devoir choisir entre comprendre et utiliser les ordinateurs<sup>801</sup> ».

### Images et sons après le numérique : mutation, glissement, rupture

En ce qui concerne l'image et le son, il nous semble que l'ère numérique ne constitue pas à proprement parler une rupture, mais plutôt une transformation, qui combine de l'ancien et du nouveau. Elle n'est pas non plus, comme le décrit Bachimont, « un progrès, ou une régression, mais une mutation qui s'impose à nous même si nous en sommes les auteurs<sup>802</sup> ». Les évolutions des relations entre image et son au fil des technologies nécessitent en conséquence d'être étudiées aussi bien dans leurs fractures, que dans leurs continuités. C'est ce que préconise la théoricienne Yvonne Spielmann pour l'image animée<sup>803</sup>, en proposant le concept de déplacement. Il est alors probablement utile de penser aussi l'image et le son, comme le fait Bidhan Jacobs, dans un sens plus large que celui du numérique. Pour lui, il s'agit de penser le « signal filmique » : « l'essence même du son et de l'image détectés ou simulés, de l'information matérielle codée qui se décline selon les matériaux des supports argentique, vidéographique et numérique, et nécessitant des dispositifs de reconstruction pour rayonner dans l'es-

798 Nicole Brenez, « Flux, flash et prisme. Vitesses critiques dans l'histoire du cinéma », *op. cit.*

799 Vilém Flusser, « Une révolution dans le domaine des images », in Vilém Flusser, *La civilisation des médias*, *op. cit.*, p. 62.

800 « En effet, où que l'on aille se promener, on y traîne avec soi des appareils émetteurs de ce flot d'images qui nous aveuglent et nous étourdissent. Ces appareils que nous trimballons partout, il n'est même plus nécessaire qu'ils pendouillent devant nos ventres. Nous les portons déjà dans nos ventres, et ils s'y déclenchent, s'y déroulent et s'y réenroulent en nous-mêmes. Nous sommes livrés à ce déferlement d'images sonorisées » (*Ibid.*, p. 57-58).

801 Olia Lialina, « L'utilisateur Turing-complet », *op. cit.*

802 Bruno Bachimont, « Signes formels et computation numérique : entre intuition et formalisme. Critique de la raison computationnelle », *op. cit.*

803 « Les débats courants portant sur l'image et la visualité abordent le déplacement dans les arts causé par les nouvelles technologies en termes de rupture et de fracture alors que les similitudes et les continuités, même dans les exposés historiques, ne sont en comparaison guère examinées » (Yvonne Spielmann, « De la temporalité à la spatialité : intermédia, hypermédia et hybridation », in François Soullages (dir.), *Dialogues sur l'art et la technologie autour d'Edmond Couchot*, Paris, L'Harmattan, 2001, p. 102-103).

pace du spectateur<sup>804</sup> ». À leurs différents âges (celui de l'optique, celui de l'analogique, celui du numérique), les images et les sons ont muté technologiquement ensemble (avec parfois certains décalages). Ils s'inscrivent désormais sur les mêmes supports, circulent grâce aux mêmes infrastructures, et sont soumis aux mêmes algorithmes.

Prospectivement, il est raisonnable de penser que les images et les sons, car ils font appel à deux sens liés et constamment en éveil — l'ouïe, la vision — continueront de dialoguer ensemble dans d'autres contextes techniques. L'ère post-numérique du quantique, ou de la bio-informatique, vont probablement contribuer à faire émerger de nouvelles formes audio-visuelles, qui cohabiteront avec des média plus anciens. Il appartient aux artistes d'en traduire les langages et d'en dévoiler les ossatures.

---

804 Bidhan Jacobs, *Esthétique du signal*, op. cit., p. 512.

## Bibliographie

### Ouvrages

- ALBERS, Josef, *L'interaction des couleurs*, Paris, Hazan, 2013.
- ALMIRON, Miguel, *À la Recherche du visage perdu dans l'art numérique*, Villeneuve-d'Ascq, Presses universitaires du Septentrion, coll. « Arts du spectacle — Images et sons », 2022.
- AUMONT, Jacques, *Du visage au cinéma*, Paris, Cahiers du cinéma, coll. « Essais », 1992.
- BALPE, Jean-Pierre, *Hyperdocuments, hypertextes, hypermédias*, Paris, Eyrolles, 1990.
- BAUDOIN, Olivier, *Pionniers de la musique numérique*, Sampzon, Delatour, coll. « Pensée musicale », 2012.
- BEAUVAIS, Yann (dir.), *Malcolm Le Grice. Le temps des images*, Dijon, Les presses du réel, 2015.
- BEAUVAIS, Yann, BOUHOURS, Jean-Michel (dir.), *Monter-sampler : l'échantillonnage généralisé*, Paris, Centre Pompidou/Scratch Projection, 2000.
- BEAUVAIS, Yann (dir.), *Musique Film*, Paris, Scratch/Cinémathèque française, 1986.
- BEUGNET, Martine, *Le cinéma et ses doubles. L'image de film à l'ère du foundfootage numérique et des écrans de poche*, Lormont, Le bord de l'eau, 2021.
- BELLOUR, Raymond, *L'Entre-images : photo, cinéma, vidéo*, Paris, La Différence, coll. « Les Essais », 2002.
- BEYAERT-GESLIN, Anne, *Sémiotique du portrait. De Dibutade au selfie*, Paris, De Boeck Supérieur, coll. « Culture & Communication », 2017.
- BOILLAT, Alain, *Du bonimenteur à la voix-over, Voix-attraction et voix-narration au cinéma*, Lausanne, Antipodes, coll. « Médias et histoire », 2007.
- BOISSIER, Jean-Louis, PINKAS, Daniel (dir.), *Jouable. Art, jeu et interactivité*, Genève, Centre pour l'image contemporaine/ Haute école d'arts appliqués HES/ École nationale supérieure des arts décoratifs/ Ciren, Université Paris 8, 2004.
- BOISSIER, Jean-Louis, *La relation comme forme. L'interactivité en art*, Genève, MAMCO, 2004.
- BOLAY, Jean-Marie, *György Kepes. Du langage visuel à l'art environnemental*, Genève, MétisPresses, 2018.
- BOSSEUR, Jean-Yves, *Musique et Arts Plastiques, Interactions au XX<sup>e</sup> siècle*, Paris, Minerve, coll. « Musique Ouverte », 2006.
- BOSSEUR, Jean-Yves, *Vocabulaire de la musique contemporaine*, Paris, Minerve, 2013.
- BOSSIS, Bruno, *La Voix et la machine. La vocalité artificielle dans la musique contemporaine*, Rennes, Presses Universitaires de Rennes, coll. « Æsthetica », 2005.

- BOURASSA, Renée, POISSANT, Louise (dir.), *Avatars, personnages et acteurs virtuels*, Québec, Presses de l'Université du Québec, coll. « Esthétique », 2013.
- BREGMAN, Albert, *Auditory Scene Analysis*, Cambridge, MIT Press, 1990.
- BRENEZ, Nicole Brenez, LEBRAT, Christian (dir.), *Jeune, Dure Et Pure ! Une Histoire Du Cinéma D'avant-Garde Et Expérimental*, Paris, Cinémathèque Française/Mazzotta, 2001.
- BRETON, Philippe, *L'éloge de la parole*, Paris, La Découverte, 2007.
- BRIDLE, James, *Un nouvel âge de ténèbres. La technologie et la fin du futur*, Paris, Allia, 2022.
- CAILLAUD, Bernard, *La création numérique visuelle, Aspects du Computer Art depuis ses origines*, Paris, Europa, 2001.
- CARVALHO, Ana, LUND, Cornelia (dir.), *The Audiovisual Breakthrough*, Berlin, autoédition, 2015.
- CASILI, Antonio A., *En attendant les robots, Enquête sur le travail du clic*, Paris, Seuil, coll. « La Couleur des idées », 2019.
- CASTANT, Alexandre, *Planètes sonores. Radiophonie, arts, cinéma*, Blou, Monografik, 2010.
- CAUQUELIN, Anne, *L'exposition de soi. Du journal intime aux Webcams*, Paris, Eshel, 2003.
- CHATEAU, Dominique (dir.), *La direction de spectateurs. Création et réception au cinéma*, Bruxelles, Les Impressions Nouvelles, coll. « Réflexions faites », 2015.
- CHIK, Caroline, *L'image paradoxale — Fixité et mouvement*, Villeneuve-d'Ascq, Presses universitaires du Septentrion, 2011.
- CHION, Michel, *La voix au cinéma*, Cahiers du Cinéma, coll. « Cinéma Essais », Paris, 1984.
- CHION, Michel, *L'audio-vision. Son et image au cinéma*, Paris, Armand Colin, 1990.
- CHION, Michel, *Le son*, Paris, Nathan, 2000.
- CHOPIN, Henri, *Poésie sonore internationale*, Paris, Jean-Michel Place, 1979.
- COUCHOT, Edmond, HILLAIRES, Norbert (dir.), *L'art numérique. Comment la technologie vient au monde de l'art*, Paris, Flammarion, 2003.
- CRICQUI, Jean-Pierre (dir.), *Christian Marclay : Replay, catalogue d'exposition*, Paris, Cité de la musique/Réunion des musées nationaux, 2007.
- DALLET, Jean-Marie, GERVAIS, Bertrand, (dir.), *Architectures de mémoire*, Dijon, Les presses du réel, 2019.
- DALLET, Jean-Marie (dir.), *Cinéma, interactivité et société*, Université de Poitiers & CNRS/VDMC, 2013.
- DANIELS, Dieter, NAUMANN, Sandra (dir.), *Audiovisuology Compendium : An Interdisciplinary Survey of Audiovisual Culture*, Köln, Walther König, coll. « See This Sound », 2010.

- DANIELS, Dieters, NAUMANN, Sandra (dir.), *Audiovisuology 2. Essays. Histories and Theories of Audiovisual Media and Art*, Köln, Walther König, coll. « See This Sound », 2011.
- DEPOCAS, Alain, IPPOLITO, Jon, JONES, Caitlin (dir.), *L'approche des médias variables : la permanence par le changement*, New York/Québec, Guggenheim Museum/Fondation Daniel Langlois, 2003.
- DRUCKER, Johanna, *Visualisation. L'interprétation modélisante*, Paris, B42, 2020.
- DULAC, Germaine, *Écrits sur le cinéma (1919-1937)*, Paris, Paris Expérimental, 1994.
- DUPLEIX, Sophie, LISTA, Marcella (dir.), *Sons et lumières. Une histoire du son dans l'art du XX<sup>e</sup> siècle*, catalogue d'exposition, Paris, Centre Pompidou, 2004.
- ECO, Umberto, *L'Œuvre ouverte*, Paris, Seuil, coll. « Points », 1965.
- EISENSTEIN, S. M., *Le film : sa forme, son sens*, Paris, Christian Bourgois, 1976.
- EPSTEIN, Jean, *L'Intelligence d'une machine*, Paris, Jacques Melot, 1946.
- ETIENNE, Yvan, SALADIN, Matthieu (dir.), *Tacet n° 3*, « De l'espace sonore », Dijon, Haute école des arts du Rhin/Les Presses du réel, coll. « OH CET ECHO », 2014.
- FATUS, Claude, *Vocabulaire des nouvelles technologies musicales*, Paris, Minerve, coll. « Musique ouverte », 1994.
- FLUSSER, Vilém, *Choses et non-choses. Esquisses phénoménologiques*, Paris, Jacqueline Chambon, 1996.
- FLUSSER, Vilém, *La civilisation des médias*, Belval, Circé, 2006.
- FOURMENTRAUX, Jean-Paul, *antiDATA, la désobéissance numérique. Art et hacktivisme technocritique*, Dijon, Les Presses du réel, coll. « Perceptions », 2020.
- FOURMENTRAUX, Jean-Paul (dir.), *Images interactives. Art contemporain, recherche et création numérique*, Bruxelles, La Lettre volée, coll. « Essais », 2017.
- FRIELING, Rudolf, DANIELS, Dieter (dir.), *Media Art Net 1 : Survey of Media Art*, Vienne/New York, Goethe-Institut/ZKM, Springer, 2004.
- FRIELING, Rudolf, DANIELS, Dieter (dir.), *Media Art Net 2 : Key Topics*, Vienne/New York, Goethe-Institut/ZKM, Springer, 2004.
- GAME, Jérôme (dir.), *Images des corps/Corps des images au cinéma*, Lyon, ENS Éditions, coll. « Signes », 2010.
- GERSTNER, Karl, *Designing Programmes. Programme as Typeface, Typography, Picture, Method*, Salestein, Niggli Verlag, 1964.
- GIBSON, Steve, ARISONA, Stefan, LEISHMAN, Donna, TANAKA, Atsu (dir.), *Live Visuals: History, Theory, Practice*, Abingdon-on-Thames, Taylor & Francis Ltd, 2022.
- GUELTON, Bernard (dir.), *Médias situés et mobilités partagées*, Paris, Hermann, 2019.
- GUNTHER, André, *L'Image partagée. La photographie numérique*, Paris, Textuel, 2015.
- GODARD, Jean-Luc, *Godard par Godard. Les Années Cahiers*, Paris, Flammarion, 1989.

- GONZALES, Rita (dir.), *Christian Marclay. Sound Stories*, catalogue d'exposition, Santa Monica, LACMA/Snap Inc., 2019.
- GOODY, Jack, *La Raison graphique, La domestication de la pensée sauvage*, Paris, Minuit, 1977.
- GREENE, Rachel, *L'art Internet*, Paris, Thames & Hudson, coll. « L'univers de l'art », 2005.
- HADJITHOMAS, Joana, JOREIGE, Khalil, *Les rumeurs du monde. Repenser la confiance à l'âge de l'Internet*, Berlin, Sternberg Press, 2015.
- HAVERKAMP, Michael, *Synesthetic Design: Handbook for a Multi-Sensory Approach*, Berlin, Walter de Gruyter, 2012.
- HAYLES, N. Katherine, *Lire et penser en milieux numériques. Attention, récits, technogène*, Grenoble, Ellug, 2016.
- HAYLES, N. Katherine, *Parole, écriture, code*, Dijon, Les presses du réel, 2015.
- ROGERS, Holly, *Sounding the Gallery: Video and the Rise of Art-Music*, Oxford, Oxford University Press, 2013.
- HEGARTY, Paul, *Rumour and Radiation. Sound in Video Art*, Londres, Bloomsbury, 2015.
- HERMANN, Thomas, HUNT, Andy, NEUHOFF, John (dir.), *The Sonification Handbook*, Berlin, Logos Publishing House, 2011.
- HEUDIN, Jean-Claude, *Les Créatures artificielles : Des automates aux mondes virtuels*, Paris, Odile Jacob, 2008.
- HEUDIN, Jean-Claude, *Comprendre le Deep Learning : une introduction aux réseaux de neurones*, Paris, Science eBook, version Kindle, 2016.
- HIGGINS, Hannah B., KAHN, Douglas (dir.), *Mainframe Experimentalism: Early Computing and the Foundations of the Digital Arts*, Berkeley, University of California Press, 2012.
- HODEIR, André, *Les formes de la musique*, Paris, PUF, 2003.
- HOFFMANN, Kay, ELSAESSER, Thomas (dir.), *Cinema Futures : Cain, Abel Or Cable ? The Screen Arts in the Digital Age*, Amsterdam, Amsterdam University Press, 1996.
- HOFMANN, Armin, *Graphic Design Manual: Principles and Practice*, New York, Van Nostrand Reinhold, 1965.
- IKEDA, Ryoji, *Continuum*, catalogue d'exposition, Paris, Xavier Barral/Centre Pompidou, 2018.
- IKEDA, Ryoji, *Datamatics*, Milan, Charta, 2012.
- JACOBS, Bidhan, *Esthétique du signal. Hacker le filmique*, Paris, Mimésis, coll. « Images, médiums », 2022.
- JULLIER, Laurent, *Les sons au cinéma et à la télévision : Précis d'analyse de la bande-son*, Paris, Armand Collin, coll. « Cinéma et audiovisuel », 1997.

- KALTENECKER, Martin, LE BAIL, Karine (dir.), *Pierre Schaeffer. Les constructions impatientes*, Paris, CNRS Éditions, 2012.
- KANDINSKY, *Point et Ligne sur plan*, Paris, Gallimard, coll. « Folio essais », 1991.
- KEPES, György, *Language of Vision*, Chicago, New York, Dover Publications, 1995.
- KITTLER, Friedrich, *Gramophone, Film, Typewriter*, Dijon, Les presses du réel, 2018.
- LANGOIS, Philippe, *Les cloches d'Atlantis : musique électroacoustique et cinéma, archéologie et histoire d'un art sonore*, Paris, Éditions MF, coll. « Répercussions », 2012.
- LARSONNEUR, Claire, REGNAULD, Arnaud, CASSOU-NOGUÈS, Pierre, TOUIZA, Sara (dir.), *Le Sujet Digital*, Dijon, Les presses du réel, 2015.
- LARTIGAUD, David-Olivier (dir.), *Art++*, Orléans, HYX, coll. « Script », 2011.
- LEAVITT, Ruth (dir.), *Artist and Computer*, New York, Harmony Books, 1976.
- LEBRAT, Christian, *Peter Kubelka*, Paris, Paris Expérimental, 1992.
- LE GRICE, Malcolm, *Abstract Film and beyond*, Cambridge, MIT Press, 1977.
- LE GRICE, Malcolm, *Experimental Cinema in the Digital Age*, Londres, BFI Publishing, 2002.
- LIMARE, Sophie, *Surveiller et sourire : les artistes visuels et le regard numérique*, Montréal, Les Presses de l'Université de Montréal, 2015.
- LISTA, Marcella (dir.), *Paul Klee, Polyphonies*, catalogue d'exposition, Paris, Actes Sud/Cité de la Musique, 2011.
- MACKENZIE, Scott, MARCHESSAULT, Janine (dir.), *Process Cinema. Handmade Film in the Digital Age*, Montréal, McGill-Queen's University Press, 2019.
- MAEDA, John, *Code de création*, Londres, Thames & Hudson, 2004.
- MAEDA, John, *De la simplicité*, Paris, Payot, 2007.
- MAEDA, John, *Design By Numbers*, Cambridge,, MIT Press, 1999.
- MANNONI, Laurent, *Le grand art de la lumière et de l'ombre. Archéologie du cinéma*, Paris, Nathan, 1994.
- MANOVICH, Lev, *Le langage des nouveaux médias*, Dijon, Les presses du réel, 2010.
- MASURE, Anthony, *Design et humanités numériques*, Paris, B42, coll. « Esthétique des données », 2017.
- MICHAUD, Philippe-Alain, BENSON, Timothy O. (dir.), *Hans Richter, la traversée du siècle*, Centre Pompidou-Metz, 2013.
- MIGAYROU, Frédéric (dir.), *De Stijl. 1917-1931*, Paris, Centre Pompidou, 2010.
- MIGAYROU, Frédéric (dir.), *Coder le monde*, Orléans, HYX/Centre Pompidou, coll. « Script », 2018.
- MOLES, Abraham A., *Art et ordinateur*, Paris, Blusson, 1990.

- MOHR, Manfred, *Manfred Mohr, Computer Graphics — Une Esthétique Programmée*, catalogue d'exposition, Paris, ARC/Musée d'Art Moderne de la Ville de Paris, 1971.
- NAESSENS, Ophélie, *Portraits parlants dans l'art vidéo. La parole vivante dans les pratiques artistiques des années soixante-dix à nos jours*, Paris, L'Harmattan, 2021.
- NARDIN, Patrick, PERRET, Catherine, PHAY, Soko, SEIDERER, Anna (dir.), *Archives au présent*, Vincennes, Presses universitaires de Vincennes, coll. « Esthétiques hors cadre », 2017.
- NEUHAUS, Max, *Évoquer l'auditif*, catalogue d'exposition, Nice, Villa Arson, 1995.
- ODIN, Paul-Emmanuel, *L'inversion temporelle au cinéma*, Dijon, Les presses du réel, coll. « Al Dante », 2014.
- OVIDE, *Les Métamorphoses*, Livre III 383-410, Édition de Jean-Pierre Néraudeau, Paris, Gallimard, coll. « Folio Classique », 1992.
- PARIKKA, Jussi, *Qu'est-ce que l'archéologie des média ?*, Saint-Martin-d'Hères, UGA Éditions, coll. « Savoir littéraires et imaginaires scientifiques », 2017.
- PARIKKA, Jussi, *Trois essais sur l'écologie des media*, Varennes, T&P Publishing, coll. « Iconodules », 2021.
- PATTERSON, Zabet, *Peripheral Vision: Bell Labs, the S-C 4020, and the Origins of Computer Art*, Cambridge, MIT Press, 2015.
- PELÉ, Gérard, *Études sur la perception auditive*, Paris, L'Harmattan, coll. « Arts & Sciences de l'Art », 2016.
- PASCUALE, Frank, *Black Box Society. Les algorithmes secrets qui contrôlent l'économie et l'information*, Limoges, FYP Éditions, coll. « Présence/Essai », 2015.
- PICARD, Rosalind, *Affective Computing*, Cambridge, MIT Press, 1997.
- PHILIZOT, Vivien (dir.), *Technique et design graphique. Outils, médias, savoirs*, Paris, B42, 2020.
- PIERCE, John R., *Le son musical. Musique, acoustique et informatique*, Paris, Belin, coll. « Pour la science », 1984.
- PISANO, Giusy, *Une archéologie du cinéma sonore*, CNRS Éditions/OpenEdition Books, 2004.
- POISSANT, Louise (dir.), *Dictionnaire des arts médiatiques*, Québec, Presses de l'Université du Québec, 1997.
- POSCHARDT, Ulf, *DJ Culture*, Paris, Éditions Kargo/L'Éclat, 2002.
- PUCKETTE, Miller, *The Theory and Technique of Electronic Music*, Singapour, World Scientific Publishing Company, 2006.
- REAS, Casey, MCWILLIAMS, Chandler (dir.), *Form+Code in Design, Art, and Architecture*, New York, Princeton Architectural Press, 2010.
- REICHARDT, Jasia (dir.), *Cybernetic Serendipity. The computer and the arts*, catalogue d'exposition, Studio International Fondation, New York/Washington, 1968-2018.

- REICHARD, Jasia, *The Computer in Art*, Londres, Studio Vista, 1971.
- RENARD Caroline (dir.), *Images numériques ?*, Aix-en-Provence, Presses universitaires de Provence, 2014.
- RICHTER, Hans, *Dada, art et anti-art*, Bruxelles, La connaissance, 1965.
- RISSET, Jean-Claude, *Écrits — Volume I. Composer le son : repères d'une exploration du monde sonore numérique*, Paris, Hermann, coll. « GREAM/Création contemporaine », 2014.
- ROADS, Curtis, *L'audio numérique. Musique et informatique*, Paris, Dunod, coll. « Audio-Photo-Vidéo », 2e édition, 2007.
- ROUFFINEAU, Gilles, *Éditions off-line : projet critique de publications numériques 1989-2001*, Paris, B42, 2018.
- RUSHKOFF, Douglas, *Les 10 commandements de l'ère numérique*, Limoges, FYP, coll. « Innovation », 2012.
- RUSSETT, Robert, STARR, Cecile (dir.), *Experimental Animation. Origins of a New Art*, Cambridge, Da Capo Press, 1976.
- RUSSOLO, Luigi, *L'art des bruits. Manifeste futuriste 1913*, Paris, Allia, 2003.
- SCHAEFFER, Pierre, *Traité des objets musicaux*, Paris, Seuil, 1966.
- SCHAFER, Raymond Murray, *Le paysage sonore, le monde comme musique*, coll. « Domaine sauvage », Marseille, WildProject, 2010.
- SCHWARTZ, Lillian F., *The Computer Artist's Handbook: Concepts, Techniques, and Applications*, New York, WW Norton & Company, 1992.
- SMIRNOV, Andrei, *Sound in Z: Experiments in Sound and Electronic Music in Early 20th-century Russia*, Köln, Walther König, 2013.
- SOLOMOS, Makis, *De la musique au son. L'émergence du son dans la musique des XX<sup>e</sup>-XXI<sup>e</sup> siècles*, Rennes, coll. « Aesthetica », Presses Universitaires de Rennes, 2013.
- STERNE, Jonathan, *Une histoire de la modernité sonore*, Paris, La Découverte/Cité de la musique, coll. « Cultures sonores », 2015.
- SOULAGES, François (dir.), *Dialogues sur l'art et la technologie autour d'Edmond Couchot*, Paris, L'Harmattan, 2001.
- SZENDY, Peter, *Sur écoute. Esthétique de l'espionnage*, Paris, Minuit, coll. « Paradoxe », 2007.
- TISSERON, Serge, *L'intimité surexposée*, Paris, Hachette Littératures, 2002.
- TKACZYK, Viktoria, MILLS, Mara, HUI Alexandra (dir.), *Testing Hearing*, Oxford, Oxford University Press, 2020.
- VAN ASSCHE, Christine, *Sonic Process. Une nouvelle géographie des sons*, catalogue d'exposition, Paris, Centre Pompidou, 2002.
- VAN DOESBURG, Theo, *Principes Fondamentaux de l'Art néo-Plastique*, Paris, ENSBA, 2003.

VANDENDORPE, Christian, *Du papyrus à l'hypertexte. Essai sur les mutations du texte et de la lecture*, Paris, La découverte, coll. « Sciences et société », 1999.

VARÈSE, Edgar, *Écrits*, Paris, Christian Bourgois, 1983.

VIAL, Stéphane, *L'être et l'écran : comment le numérique change la perception*, Paris, PUF, coll. « Hors collection », 2013.

WARHOL, Andy, *Entretiens 1962-1987*, Paris, Grasset, 2005.

WEISSBERG, Jean-Louis, *Présences à distance. Déplacement virtuel et réseaux numériques : pourquoi nous ne croyons plus à la télévision*, Paris, L'Harmattan, 1999.

WHITNEY, John, *Digital Harmony. On the Complementarity of Music and Visual Art*, Peterborough, McGraw-Hill, 1980.

WILLOUGHBY, Dominique, *Le cinéma graphique. Une histoire des dessins animés : des jouets d'optique au cinéma numérique*, Paris, Textuel, 2009.

XENAKIS, Iannis, *Musiques formelles*, Revue Musicale n° 253-254, Paris, Stock, 1963.

YOUNG Paul, DUNCAN, Paul (dir.), *Le cinéma expérimental*, Paris, Taschen, 2009.

YOUNGBLOOD, Gene, *Expanded Cinema*, Dutton, New York, 1970.

ZILIO, Marion, *Faceworld. Le visage au XXI<sup>e</sup> siècle*, Paris, PUF, coll. « Perspectives critiques », 2018.

ZÉNOUDA, Hervé, *Les images et les sons dans les hypermédias artistiques contemporains*, Paris, L'Harmattan, 2004.

ZUBOFF, Shoshana, *L'âge du capitalisme de surveillance*, Honfleur, Zulma, coll. « Essais », 2022.

## Ouvrages collectifs et revues

MCD — Magazine des Cultures Digitales #66, hors-série, « Machines d'écritures », mars/mai 2012.

MCD — Magazine des Cultures Digitales, hors-série, « Live A/V. Performances audiovisuelles », avril 2010.

*Laboratoires de l'art*, catalogue d'exposition, Luxembourg, Paris, Mudam Luxembourg, Hermann, Musée des arts et métiers-le Cnam, 2016.

*Piotr Kowalski. Time Machine et projets*, catalogue d'exposition, Paris, Centre Georges Pompidou, Musée national d'art moderne, 1981.

## Thèses

ABENAVOLI, Lorella, *Le son plastique : empreindre le flux et l'inouï. Sonification et audification dans l'art de l'installation*, sous la direction de Louise Poissant, Université du Québec à Montréal, 2017.

CHARLES, Carlos Lopez, *La convergence musicale entre les médias visuels et musicaux dans la création de la musique visuelle*, sous la direction de Anne Sédès, École doctorale Esthétique, sciences et technologie des arts, Université de Paris 8, Saint-Denis, 2015.

CICILIATO, Vincent, *Glitch(s) imago-sonore(s) : du processus d'objectivation par découpe et synchronisation son-image à une esthétique du micro-mouvement*, sous la direction de Françoise Parfait, Amiens, 2010.

KRATKY, Andreas, *Les auras numériques : pour une poétique de la base de données*, thèse en arts plastiques et sciences de l'art, sous la direction d'Anne-Marie Duguet, Paris, Université Paris 1 Panthéon-Sorbonne, 2013

## Articles

FANO, Michel, « Musique et film : filmusique », In *Vibrations*, n° 4, « Les musiques des films », 1987, p. 27.

SHANNON, Claude, « A Mathematical Theory of Communication », in *The Bell System Technical Journal* n° 27, New York, AT&T, 1948.

SEKULER, Robert, SEKULER, Allison B., LAU, Renee, « Sound alters visual motion perception », *Nature* n° 385, p. 308, janvier 1997.

SCHAEFFER, Pierre, « Le contrepoint du son et de l'image », in *Cahiers du cinéma*, n° 108, juin 1960.

TURING, Alan, « On computable numbers, with an application to the Entscheidungsproblem », in *Proceedings of the London Mathematical Society*, vol. 42, n° 2, Londres, 1936.

VARGOZ, Frédérique, GUEZ, Emmanuel, « Une histoire de l'ordinateur du point de vue de la théorie des média » in *Cahiers philosophiques* n° 141, « La révolution informatique », Paris, Canopé, mai 2015, p. 62.

## Sitographie

### Articles

AIMÉ, Xavier, « Intelligence artificielle et psychiatrie : noces d'or entre Eliza et Parry », in L'information psychiatrique n° 93, 2017, p. 51-56, <https://www-cairn-info.ezpaarse.univ-paris1.fr/revue-l-information-psychiatrique-2017-1-page-51.htm>, consulté le 30/01/23.

ANGELL, Callie, « Doubling the Screen: Andy Warhol's Outer and Inner Space », Millennium Film Journal No. 38, 2002, <http://www.mfj-online.org/journalPages/MFJ38/angell.html>, consulté le 30/01/23.

AUCOUTURIER, Jean-Julien, GUEROUAOU, Nadia, VAIVA, Guillaume, « L'étonnante acceptabilité des deep-fake », Le journal CNRS, janvier 2022, <https://lejournel.cnrs.fr/billets/letonnante-acceptabilite-des-deep-fake>, consulté le 30/01/23.

BACHIMONT, Bruno, « Signes formels et computation numérique : entre intuition et formalisme. Critique de la raison computationnelle », 2004, [http://www.utc.fr/~bachimon/Publications\\_attachments/Bachimont.pdf](http://www.utc.fr/~bachimon/Publications_attachments/Bachimont.pdf), consulté le 30/01/23.

BACHIMONT, Bruno, « L'archive et la massification des données : une nouvelle raison numérique », Gazette des archives n° 245, 2017, [https://www.persee.fr/issue/gazar\\_0016-5522\\_2017\\_num\\_245\\_1?sectionId=gazar\\_0016-5522\\_2017\\_num\\_245\\_1\\_5510](https://www.persee.fr/issue/gazar_0016-5522_2017_num_245_1?sectionId=gazar_0016-5522_2017_num_245_1_5510), consulté le 30/01/23.

BAUDOUIN, Philippe, « Du Hörspiel », revue Syntone, 2011, <http://syntone.fr/du-horspiel/>, consulté le 30/01/23.

BEAUSSE, Pascal, « Candice Breitz : images infectées », texte de présentation de la soirée Candice Breitz, jeudi 28 avril 2005, Palais de Tokyo, <https://www.pointligneplan.com/document/candice-breitz-images-infectees/>, consulté le 30/01/23.

BEYLS, Peter, « WindChime: database driven mapping Of Complex dynamical systems », 2014, <http://2014.xcoax.org/pdf/xcoax2014-Beyls.pdf>, consulté le 30/01/23.

BOISCLAIR, Louise, « David Rokeby : Je suis un artiste interactif et je construis des expériences », Archée, revue d'art en ligne : arts médiatiques & cyberculture, 2011, <http://archiee.qc.ca/ar.php?page=article&no=383>, consulté le 30/01/23.

BOISSIER, Jean-Louis, « Masaki Fujihata, "Le linéaire actif". Remarques sur les Field-Works de Masaki Fujihata », [http://www.arpla.fr/canal20/adnm/wp-pdf/Fujihata\\_lineaire\\_actif.pdf](http://www.arpla.fr/canal20/adnm/wp-pdf/Fujihata_lineaire_actif.pdf), consulté le 30/01/23.

BOISSIER, Jean-Louis, DALLET, Jean-Marie, entretien avec Piotr Kowalski, Montrouge, le 24 septembre 1992, [http://www.ciren.org/artifice/artifices\\_2/kowalski.html](http://www.ciren.org/artifice/artifices_2/kowalski.html), consulté le 30/01/23.

- BONIN, Vincent, « Digital Image Articulator », Fondation Langlois, 2004, <https://www.fondation-langlois.org/html/f/page.php?NumPage=457>, consulté le 30/01/23.
- BOOTZ, Philippe, « Quelle est l'histoire de l'animation numérique ? », Leonardo/Olats, 2006, [http://archive.olats.org/livresetudes/basiques/animationnumerique/3\\_basiquesAN.php](http://archive.olats.org/livresetudes/basiques/animationnumerique/3_basiquesAN.php), consulté le 30/01/23.
- BOUCHARD, Vincent, « Du Nagra au caméscope : questions de synchronisation image/son », Intermédialités n° 19, printemps 2012, « Synchroniser », 2012, <https://www.erudit.org/en/journals/im/2012-n19-im0305/1012659ar/>, consulté le 30/01/23.
- BRENEZ, Nicole, « Flux, flash et prisme. Vitesses critiques dans l'histoire du cinéma », in Le Genre humain n° 49, 2010, <https://www-cairn-info.ezpaarse.univ-paris1.fr/revue-le-genre-humain-2010-1-page-125.htm>, consulté le 30/01/23.
- CAUX, Jacqueline, « Les violons de Laurie Anderson », in Art Press n° 285, 2002.
- CASILLI, Antonio, « L'avènement du dresseur d'intelligences artificielles », Le Monde, octobre 2011, [https://www.lemonde.fr/economie/article/2020/02/06/qui-travaillera-demain-les-robots-ou-les-humains\\_6028592\\_3234.html](https://www.lemonde.fr/economie/article/2020/02/06/qui-travaillera-demain-les-robots-ou-les-humains_6028592_3234.html), consulté le 30/01/23.
- CASILLI, Antonio, « Les humains derrière Cortana », dossier « Derrière les assistants vocaux, des humains vous entendent », La Quadrature du Net, mai 2018, [https://www.laquadrature.net/2018/05/18/temoin\\_cortana/](https://www.laquadrature.net/2018/05/18/temoin_cortana/), consulté le 30/01/23.
- CHION, Michel, « Glossaire », entrée « Acousmatisation », <http://www.lampe-tempete.fr/ChionGlossaire.html>, consulté le 30/01/23.
- CLAUSS, Nicolas, Rencontre avec Xavier Thomas, Radio Grenouille, extrait, novembre 2011, <http://www.nicolasclauss.com/terresarbitraires/TA.pdf>, consulté le 30/01/23.
- COLIN, Cécile, RADEAU, Monique, « Les illusions McGurk dans la parole : 25 ans de recherches », 2003, [https://www.persee.fr/doc/psy\\_0003-5033\\_2003\\_num\\_103\\_3\\_29649](https://www.persee.fr/doc/psy_0003-5033_2003_num_103_3_29649), consulté le 30/01/23.
- COUPRIE, Pierre, « Les théories de Pierre Schaeffer », <http://www.olats.org/pionniers/pp/schaeffer/theorieSchaeffer.php>, OLATS, 2000, consulté le 30/01/23.
- COURCHESNE, Luc, PERRON, Bernard, « T'es où [Luc Courchesne] ? », Intermédialités/Intermediality n° 9, p. 157–175, 2007, <https://doi.org/10.7202/1005535ar>, consulté le 30/01/23.
- COX, Geoff, MCLEAN, Alex, WARD, Adrian, « L'esthétique du code génératif » (The Aesthetics of Generative Code), 2001, traduit dans le cadre de l'ouvrage *Art++* publié aux Éditions HXX, en juin 2011, sous la direction de David-Olivier Lartigaud, [https://monoskop.org/images/a/ae/Cox\\_McLean\\_Ward\\_2000\\_2011\\_L\\_esthetique\\_du\\_code\\_generatif.pdf](https://monoskop.org/images/a/ae/Cox_McLean_Ward_2000_2011_L_esthetique_du_code_generatif.pdf), consulté le 30/01/23.
- DAGENAIS, Francine, « Programmation : l'impact de la programmation, l'impact de la programmation », in Revue Intermédialités/Intermediality, n° 13, printemps 2009, p. 17–37,

<https://www.erudit.org/en/journals/im/2009-n13-im3872/044038ar/>, consulté le 30/01/23.

DALLET, Jean-Marie, glossaire du rapport d'activité du programme Figures de l'inter-activité, [http://www.dallet.net/PAGEStheorie/figures\\_DAP/figures\\_DAP\\_1/B\\_r\\_DAP\\_7.html](http://www.dallet.net/PAGEStheorie/figures_DAP/figures_DAP_1/B_r_DAP_7.html), consulté le 30/01/23.

DANIELS, Dieter, « Sound & Vision in Avantgarde & Mainstream », 2004, [http://www.medienkunstnetz.de/themes/image-sound\\_relations/sound\\_vision/13/](http://www.medienkunstnetz.de/themes/image-sound_relations/sound_vision/13/), consulté le 30/01/23.

DEMAN, Samantha, « La mécanique sensible d'Eric Vernhes », ArtsHebdoMédias, 2016, <https://www.artshebdomedias.com/article/mecanique-sensible-deric-vernhes/>, consulté le 30/01/23.

ERNST, Wolfgang, « Beyond the Archive: Bit Mapping », Media Art Net, 2004, [http://www.medienkunstnetz.de/themes/mapping\\_and\\_text/beyond-the-archive/1/](http://www.medienkunstnetz.de/themes/mapping_and_text/beyond-the-archive/1/), consulté le 30/01/23.

ESS, Xavier, Interview de David Rokeby pour l'émission Cybercafé21 (RTBF), au Théâtre Mercelis, Bruxelles, décembre 2003, <https://legacy.imal.org/drokeby/archives/interviewfr.html>, consulté le 30/01/23.

FANEN, Sophian, Dossier « Alexa, l'espionne qui m'aimait. Le web devient vocal, mais qui nous parle ? », 2018, <https://lesjours.fr/obsessions/alexamazon-espionne/>, consulté le 30/01/23.

FARGIER, Noémie, descriptif de la journée d'études « Mappemondes sonores au XXIème siècle. Une autre approche des lieux », Centre de recherches sur les arts et le langage — CRAL, <https://www.ehess.fr/fr/appel-communication/mappemondes-sonores-xxi%C3%A8me-si%C3%A8cle-autre-approche-lieux>, consulté le 30/01/23.

FOURMENTRAUX, Jean-Paul, « Net art », in Communications n° 88, 2011, <https://www-cairn-info.ezpaarse.univ-paris1.fr/revue-communications-2011-1-page-113.htm>, consulté le 30/01/23.

GAGNON, Jean, « Lynn Hershman », Fondation Langlois, 2006, <https://www.fondation-langlois.org/html/f/page.php?NumPage=168>, consulté le 30/01/23.

GARCIA, Chris, « Oral History of Lillian F. Schwartz », août 2013, <http://archive.computerhistory.org/resources/access/text/2016/06/102746737-05-01-acc.pdf>, consulté le 30/01/23.

GUEROUAOU, Nadia, « Rendre sa voix plus souriante : deepfakes et filtres vocaux émotionnels », AOC, juillet 2022, <https://aoc.media/analyse/2022/07/05/rendre-sa-voix-plus-souriante-deepfakes-et-filtres-vocaux-emotionnels/>, consulté le 30/01/23.

HERSHMAN, Lynn, « DiNA – The People's Choice », Media Art Net, 2004, <http://www.medienkunstnetz.de/source-text/150/>, consulté le 30/01/23.

JAHJAH, Marc, « Digital et Book Studies : les formes de la matérialité numérique », 2015, <https://marcjahjah.net/198-digital-book-studies-22-formes-de-materialite-numerique>, consulté le 30/01/23.

- JULLIER, Laurent, PÉQUIGNOT, Julien, « L'effet-clip au cinéma », Kinephanos, 2013, <https://www.kinephanos.ca/2013/effet-clip/>, consulté le 30/01/23.
- LANGLOIS, Philippe, « Oskar Fischinger », OLATS, 2001, <http://www.olats.org/pionniers/pp/fischinger/fischinger.php>, consulté le 30/01/23.
- LIALINA, Olia, « L'utilisateur Turing-complet », trad. Jean-François Caro, 2013, <http://reader.lgru.net/texts/lutilisateur-turing-complet/>, consulté le 30/01/23.
- MANOVICH, Lev, « Data Visualization as New Abstraction and as Anti-Sublime », 2002, <http://manovich.net/index.php/projects/data-visualisation-as-new-abstraction-and-anti-sublime>, consulté le 30/01/23.
- MAHLER, Louise, Interview de Corinne Castel, <http://www.ligne16.net/le-109-en-action-pour-le-festival-movimenta/>, consulté le 30/01/23.
- MORELLET, François, « Pour une peinture expérimentale programmée », article publié dans la plaquette du Groupe de recherche d'art visuel, Paris, Galerie Denise René et le GRAV, 1962, <https://francois-morellet.fr/2021/06/01/pour-une-peinture-experimentale-programmee-1962/>, consulté le 30/01/23.
- PANDILOVSKI, Melentie, Interview de David Rokeby, 2005, <https://www.videopool.org/watch-interview-with-david-rokeby/>, consulté le 30/01/23.
- PERRON, Jacques, « Lorella Abenavoli. Le Souffle de la terre », Fondation Langlois, 2000, <https://www.fondation-langlois.org/html/f/page.php?NumPage=268>, consulté le 30/01/23.
- QUENAULT, Grégoire, « Histoires et mythes de l'art vidéo. Retour sur une genèse », in Revue Ligeia 2014/2 (N° 133-136), p. 161-170, <https://www-cairn-info.ezpaarse.univ-paris1.fr/revue-ligeia-2014-2-page-161.htm>, consulté le 30/01/23.
- RAABE, Juliette, « Vidéodisque et édition », in Jacques Anis, Jean-Louis Lebrave (dir.), revue Linx, n° 17, « Le texte et l'ordinateur », 1987, [https://www.persee.fr/issue/linx\\_0246-8743\\_1987\\_num\\_17\\_1](https://www.persee.fr/issue/linx_0246-8743_1987_num_17_1), consulté le 30/01/23.
- RUTKOFF, Rebekah, « Painting by numbers: the art of Lillian Schwartz », Art Forum, octobre 2016, <https://www.artforum.com/print/201608/painting-by-numbers-the-art-of-lillian-schwartz-63453>, consulté le 30/01/23.
- SAVATON, Pierre, « Le temps fondateur de la géologie », in Romantisme n° 174, 2016, p. 39, <https://www-cairn-info.ezpaarse.univ-paris1.fr/revue-romantisme-2016-4-page-29.htm>, consulté le 30/01/23.
- SCHMITT, Antoine, « Tempest : cosmogonie d'un spectacle », <https://www.antoineschmitt.com/tempest-cosmogonie-dun-spectacle/>, consulté le 30/01/23.
- SCHNEIDER, Corinne, « Mot du jour n° 102 : Hoquet », émission Le Mot du Jour, France Musique, Mardi 17 janvier 2017, <https://www.francemusique.fr/emissions/le-mot-du-jour/le-mot-du-jour-du-mardi-17-janvier-2017-31117>, consulté le 30/01/23.

SIRON, Jacques, « Enchevêtrement de voies », Cahiers d'ethnomusicologie, 1993, <http://journals.openedition.org.ezpaarse.univ-paris1.fr/ethnomusicologie/1431>, consulté le 30/01/23.

SORREL, Vincent, « L'invention de la caméra Éclair 16 : du direct au synchrone », Revue 1895 n° 82, 2017, <https://journals-openedition-org.ezpaarse.univ-paris1.fr/1895/5330>, consulté le 30/01/23.

SPIELMANN, Yvonne, « Entretien avec Steina », Fondation Langlois, 2004, <https://www.fondation-langlois.org/html/f/page.php?NumPage=416>, consulté le 30/01/23.

SPIELMANN, Yvonne, « Steina, Violin Power, 1978 », Fondation Langlois, 2004, <https://www.fondation-langlois.org/html/f/page.php?NumPage=485>, consulté le 30/01/23.

SPIELMANN, Yvonne, « Steina et Woody Vasulka, Matrix, 1970-1972 », Fondation Langlois, 2004, <https://www.fondation-langlois.org/html/f/page.php?NumPage=495>, consulté le 30/01/23

SPIELMANN, Yvonne, « Mapping Scapes of Sound and Vision. Notes on the field work Simultaneous Echoes, an audiovisual installation », 2009, [http://www.fujihata.jp/s\\_echoes\\_09/ISEA\\_2009\\_Yvonne.pdf](http://www.fujihata.jp/s_echoes_09/ISEA_2009_Yvonne.pdf), consulté le 30/01/23.

STANLEY, Douglas Edric, « Lexicon/gps », <https://abstractmachine.net/en/lexicon/gps>, consulté le 30/01/23.

STEHURA, John, « Home Page for a Cybernetic Cinema », <https://cyberanimation.tripod.com/historypage.htm>, consulté le 30/01/23.

STIMSON, Blake, « Out in Public: Natalie Bookchin in Conversation with Blake Stimson », rhizome.org, mars 2011, <https://rhizome.org/editorial/2011/mar/09/out-public-natalie-bookchin-conversation-blake-sti/>, consulté le 30/01/23.

VARGOZ, Frédérique, « Friedrich Kittler. Y a-t-il quelqu'un dans la machine ? », MCD #66, « Machines d'écritures », mars/mai 2012, <https://www.digitalmcd.com/friedrich-kittler/>, consulté le 30/01/23.

VINTERBERG, Thomas, VON TRIER, Lars, « Dogma 95 - Le manifeste », Le Monde diplomatique, « Manière de voir » n° 151, février-mars 2017, <https://www.monde-diplomatique.fr/mav/151/VINTERBERG/57082>, consulté le 30/01/23.

VYAS, Pooja, Entretien avec Natalie Bookchin, par Zoom, février 2021, <https://sites.google.com/ucsc.edu/femexfilmarchive/filmmaker-index/natalie-bookchin>, consulté le 30/01/23.

WEINBREN, Grahame, « Time Artists: John Cage, Hollis Frampton, Pat O'Neill », 2000, [http://grahameweinbren.com/GW\\_Papers/ICCEssay.html](http://grahameweinbren.com/GW_Papers/ICCEssay.html), consulté le 30/01/23.

WEINBREN, Grahame, « The Erl King and its Digital Emulation », 2011, [http://grahameweinbren.com/GW\\_Papers/Emulating%20the%20ErlKing.html](http://grahameweinbren.com/GW_Papers/Emulating%20the%20ErlKing.html), consulté le 30/01/23.

ZÉNOUDA, Hervé, « Nouvelles organologies du sonore : les audio-games. La question des interfaces a-musicologiques », Interfaces numériques Vol. 1, n° 3, 2012, <https://www.unilim.fr/interfaces-numeriques/2456>, consulté le 30/01/23.

## Conférences et documentaires

ARCANGEL, Cory, conférence, Smithsonian American Art Museum, YouTube, 2009, <https://www.youtube.com/watch?v=oDDCoa71-yY>, consulté le 30/01/23.

CHRISTOFFEL, David, « La voix des GAFA », intervention dans le cadre de la journée d'étude « Vox Machines », organisée sous la responsabilité scientifique de Julien Drochon et Anthony Masure par le Pôle Nouveaux Médias de l'École supérieure d'art des Pyrénées (ESA, Pau-Tarbes) et le Master Design Transdisciplinaire, Cultures et Territoires, Université Toulouse Jean-Jaurès, 10 décembre 2018, <https://www.canal-u.tv/chaines/universite-toulouse-jean-jaures/vox-machines/la-voix-des-gafa-david-christoffel>, consulté le 30/01/23.

FRANCEINFO, « Terres rares : des minéraux uniques au cœur d'une bataille mondiale », Franceinfo, 02/12/21, [https://www.francetvinfo.fr/monde/terres-rares-des-mineraux-uniques-au-coeur-d-une-bataille-mondiale\\_4867259.html](https://www.francetvinfo.fr/monde/terres-rares-des-mineraux-uniques-au-coeur-d-une-bataille-mondiale_4867259.html), consulté le 30/01/23.

JOUANNAIS, Jean-Yves, *L'eau*, documentaire, 2014, <https://vimeo.com/58473589>, consulté le 30/01/23.

MEICHLER, Gisèle, MEICHLER, Luc, *Deux temps, trois mouvements*, documentaire, 1982, partie 1, <https://www.dailymotion.com/video/x411g8> et partie 2, <https://www.dailymotion.com/video/x411g8>, consulté le 30/01/23.

ONF, *À la pointe de la plume*, documentaire, 1951, [https://www.onf.ca/film/a\\_la\\_pointe\\_de\\_la\\_plume/](https://www.onf.ca/film/a_la_pointe_de_la_plume/), consulté le 30/01/23.

STEYERL, Hito, PAGLEN, Trevor, « Les images comme "Social Projectors" », conférence, Centre Pompidou, 10 novembre 2018, <https://www.centrepompidou.fr/fr/ressources/media/10fUnFE>, consulté le 30/01/23.

THISTLETHWAITE, Tom, *Earthworks by Semiconductor*, film documentaire pour la Fabrica Brighton, 2020, <https://www.youtube.com/watch?v=xkg38zms44Q&t=255s>, consulté le 30/01/23.

## Index nominum

Abenavoli, Lorella.....	112, 119	Breitz, Candice.....	284-287, 342
Abramović, Marina.....	259, 260	Brenez, Nicole.....	135, 346
Albers, Josef.....	298	Breton, Philippe.....	229
Alexïeff, Alexandre.....	27	Bridle, James.....	246, 247
Almiron, Miguel.....	226	Burroughs, William.....	132
Anderson, Laurie.....	155, 156, 250	Bute, Mary Ellen.....	41
Arcangel, Cory.....	98, 99, 196, 342	Cabaret Voltaire.....	155
Arcimboldo, Giuseppe.....	22	Cage, John.....	10, 108, 113, 172
Arnheim, Rudolf.....	303	Caillaud, Bernard.....	19, 40
Arnold, Martin.....	150, 158-160, 163, 172, 286	Carvalho, Ana.....	11
Atken, Memo.....	254	Casilli, Antonio.....	242
Aumont, Jacques.....	222, 257	Castel, Abbé Louis-Bertrand.....	22, 140
Babbage, Charles.....	43	Cerf, Vint.....	229
Bachimont, Bruno.....	12, 13, 102, 166, 167, 346	Charles, Carlos Lopez.....	26
Bailly, Jean-Christophe.....	133	Chik, Caroline.....	139
Balpe, Jean-Pierre.....	167	Chion, Michel.....	8, 139, 140, 164, 174, 220, 221, 286
Baudouin, Olivier.....	60	Chopin, Henri.....	132
Bauhaus.....	18, 23, 300, 326	Ciciliato, Vincent.....	138
Beatles.....	128	Ciocchi, Jacob.....	99
Beausse, Pascal.....	286	Citron, Jack.....	54-56
Beck, Stephen.....	72, 73, 100	Clauss, Nicolas.....	289-291
Bell Labs.....	45, 59, 60, 62, 74, 232, 300, 301, 323	Colby, Kenneth.....	229
Bentham, Jeremy.....	239	Coldcut.....	155
Bergman, Ingmar.....	223, 224	Computer Arts Society .....	47
Berio, Luciano.....	68	Computer Graphics Research Group - CGRG.....	47
Bertillon, Alphonse.....	246, 247	Computer Technique Group – CTG.....	47
Beugnet, Martine.....	188	Couchot, Edmond.....	48, 70
Beyaert-Geslin, Anne.....	341	Courchesne, Luc.....	17, 224-227, 237
Beyls, Peter.....	16, 113, 119-123, 343	Cox, Geoff.....	89
Bianchini, Samuel.....	141, 142, 144, 168, 273	Crawford, Kate.....	241, 243
Boisrobert, Anouck.....	333	Cristofol, Jean.....	14
Boissier, Jean-Louis.....	125, 143, 148, 170, 173, 209	Cros, Charles.....	126
Bonardi, Alain.....	86	Csuri, Chuck.....	56
Bookchin, Natalie... ..	17, 200, 201, 279-284, 286, 287, 342	Cuba, Larry.....	58
Bosseur, Jean-Yves.....	22	Dallet, Jean-Marie.....	145, 165, 167-169, 185, 186, 190, 340
Bossis, Bruno.....	230	Daniels, Dieter.....	73
Botticelli, Sandro.....	215	De Stijl.....	25
Boulez, Pierre.....	68	Delecluse, Christian.....	333
Brahms, Johannes.....	28	Demenÿ, Georges.....	219, 220
Bregman, Albert.....	303	Dewey-Hagborg, Heather.....	259

Dockstader, Tod.....	50, 53	Gordon, Douglas.....	136
Dodge, Charles.....	112, 232	Graeff, Werner.....	334
Doesburg, Theo van .....	300	Granular Synthesis.....	267, 269, 271, 272, 342
Dombois, Florian.....	119	Grisey, Gérard.....	9
Dukas, Paul.....	319	Grond, Florian .....	51, 111
Dulac, Germaine.....	220, 221	Guez, Emmanuel.....	13, 14, 124, 134
During, Elie.....	135	Gutmann, Newton.....	45
Duvet Brothers.....	155	Gysin, Brion.....	132
Eco, Umberto.....	67	Hadjithomas, Joana.....	292, 293, 342
Edison, Thomas A.....	126, 127, 134, 146, 201, 219, 235	Hayles, N. Katherine.....	12, 13, 167, 168, 184, 341
Eggeling, Viking.....	24-26, 33, 64, 91, 161, 299, 334	Heidsieck, Bernard.....	132
Ehrenfels, Christian von.....	302	Helmholtz, Hermann von.....	231
Ekman, Paul.....	248	Henke, Robert.....	94-98
Emergency Broadcast Network.....	155	Hentschläger, Kurt.....	267
Eno, Brian.....	75	Herndon, Holly.....	248
Epplay, Vincent.....	92, 93	Hershman Leeson, Lynn.....	233-237
Ernst, Wolfgang.....	103, 233	Heudin, Jean-Claude.....	228, 255, 256
Esling, Philippe.....	257	Hillaire, Norbert.....	48, 70
Fano, Michel.....	66, 87	Hiller, Lejaren.....	44
Fargier, Noémie.....	207	Hirsh, Hy.....	41
Fehr, Moritz.....	17, 213-216, 343	Hitchcock, Alfred.....	136, 187, 189
Fischinger, Oskar.....	27-33, 64, 91, 316, 317, 319, 320	Hoérée, Arthur.....	128
Flash, Grandmaster.....	154	Hofmann, Armin.....	91, 300, 301, 315, 316
Flusser, Vilém.....	12, 341, 346	Hohl, Frank.....	66
Fourmentaux, Jean-Paul.....	8, 249	Hollerith, Herman.....	43
Fournier, Thierry.....	206	Honegger, Arthur.....	128
Fox, Robin.....	41	Hui, Alexandra.....	108
Frampton, Hollis.....	149, 150	Huitric, Hervé.....	262-267
Franke, Herbert W.....	41	Huygues, Pierre-Damien.....	14
Friedman, Roberta.....	170-172, 175, 176	IBM.....	44, 45, 52, 54-56
Friesen, Wallace.....	248	Ikeda, Ryoji.....	16, 42, 97, 104-109, 298, 299, 341, 342
Fritz, Darko.....	20, 223	Isaacson, Leonard.....	44
Fry, Roger.....	21	Jacobs, Bidhan.....	77, 346
Fujihata, Masaki.....	17, 209-211, 343	Jacquard, Joseph Marie.....	43
Galanter, Philip.....	89	Jaubert, Maurice.....	128
Galton, Francis.....	247	Jepson, Warner.....	72, 73
Gehr, Ernie.....	149, 150	Jóhannsson, Jóhann.....	179
Gerstner, Karl.....	68	Joler, Vladan.....	241, 243
Gervais, Bertrand.....	165, 167	Joreige, Khalil.....	292, 293, 342
Godard, Jean-Luc.....	223, 224	Joy, Jérôme.....	59, 202-204, 207, 264, 342
Goethe, Johann Wolfgang von.....	171	Julesz, Béla.....	47
Goni, Kyriaki.....	238-240, 242	Jullier, Laurent.....	28, 221
Goody, Jack.....	166	Kahn, Douglas.....	10

Kandinsky, Vassily.....	23, 300, 312, 313, 315, 316	Manning, Chelsea.....	259
Kempelen, Wolfgang Von.....	231	Manovich, Lev. .16, 43, 102, 103, 106, 164-166, 177-183, 189, 199, 202	
Kepes, György.....	300	Marclay, Christian.....	17, 191-198, 242, 342, 343
Kingsley, Gershon.....	63	Marey, Étienne-Jules.....	219
Kircher, Athanasius.....	239	Marichelle, Hector.....	219
Kirsanoff, Dimitri.....	128	Martinville, Léon-Scott de.....	126
Kittler, Friedrich.....	14, 70, 124, 126, 134	Masure, Anthony.....	167
Klee, Paul.....	21, 23	Mathews, Max.....	45, 60, 63, 66, 74, 156, 232, 301
Knowlton, Ken.....	47, 59-61, 63, 65, 305	McCarthy, John.....	236
Koffka, Kurt.....	302	McGuire, Anne.....	149-151
Köhler, Wolfgang.....	302	McGurk, Harry.....	221, 303
Kowalski, Piotr.....	16, 129-134, 136, 151, 278	McLaren, Norman.....	41, 143, 330, 334
Kramer, Gregory.....	111	McLean, Alex.....	89
Kratky, Andreas.....	165, 183, 184	Mical, Abbé.....	122, 230, 231
Kronos Quartet.....	250, 251	Mignonneau, Laurent.....	333
Kubelka, Peter.....	91, 109, 323	Mills, Mara.....	108
Kubisch, Christina.....	215	Mohr, Manfred.....	47, 56, 68, 69, 91, 98, 299
Kubrick, Stanley.....	232	Mondrian, Piet.....	181
Kuntzel, Thierry.....	16, 136-138, 293	Moore, Richard.....	63, 66
La Monte Young.....	108	Morellet, François.....	67
Laboratoires Bell.....	15, 20, 45, 47	Moritz, William.....	17, 35, 37, 213-215, 343
Laënnec, René.....	110	Morris, Robert.....	215
Langheinrich, Ulf.....	267	Naessens, Ophélie.....	217
Langlois, Philippe.....	32, 75, 79, 127, 135	Nahas, Monique.....	262-267
Laposky, Ben F.....	40	Nake, Frieder.....	46, 47, 56
Lattanzi, Barbara.....	147-154	Nauman, Bruce.....	222
Laurel, Brenda.....	236	Nees, Georg.....	47, 56
Le Grice, Malcolm.....	61, 69, 70, 74-77, 79, 102, 165	Neuhaus, Max.....	10, 294
Levin, Golan.....	16, 84-88, 304, 342	Nicholson, Jack.....	284, 285
Lewin, Frank.....	63	Nicolai, Carsten.....	42
Lewis, George.....	179, 180	Noll, A. Michael.....	47, 49
Lialina, Olia.....	345, 346	Noll, Michael.....	47, 49
Liebowitz, René.....	33	Odin, Paul-Emmanuel.....	127
Lissajous, Antoine.....	41	Oram, Daphné.....	203, 210, 330, 331
Lista, Marcella.....	38, 248	Paglen, Trevor.....	237, 247, 248, 250-252, 342, 343
Locus Sonus.....	207	Paik, Nam June.....	72, 146, 147, 287
Lovelace, Ada.....	43	Paper Rad.....	98, 99
Lozano-Hemmer, Rafael.....	271, 272	Parikka, Jussi.....	124
Lumière, Frères.....	127, 223	Parke, Claire.....	61
Lund, Cornelia.....	11	Pawlowski, Auguste.....	27
Lyons, Franck.....	212	Perich, Tristan.....	323
MacDonald, John.....	221	Pfenninger, Rudolf.....	30-32, 36
Maeda, John.....	60, 85, 300		

Picard, Rosalind.....	249	Siron, Jacques.....	283
Pierce, John.....	232, 301, 316	SLIDERS_Lab.....	17, 185-190, 204, 342
Pitaru, Amit.....	330, 331	Snow, Michael.....	136
Popper, Frank.....	40	Solomos, Makis.....	9
Posters, Bill.....	259, 260	Sommerer Christa,.....	333
Pousseur, Henri.....	68	Spiegel, Laurie.....	73, 74, 78
Puckette, Miller.....	300	Spielmann, Yvonne.....	101, 346
Reed, Jerry.....	81, 83	Stanley, Douglas E.....	70, 210, 232
Reed, Lou.....	291	Stehura, John.....	20, 50-53, 297, 304
Reich, Steve.....	47, 250	Steina.....	72, 100, 156-158, 254, 270, 299, 323
Reichard, Jasia.....	46	Steyerl, Hito.....	237, 240, 248, 342, 343
Renard, Maurice.....	161	Stockhausen, Karlheinz.....	68, 132
Resonanz (Ensemble).....	253	Streep, Meryl.....	284, 286
Richter, Hans.....	24-26, 33, 91, 299	Swinton, Tilda.....	235
Rigaud, Louis.....	333	Szendy, Peter.....	239, 294
Riley, Terry.....	57, 250	Takeya, Akemi.....	268
Ringley, Jennifer.....	200	Tardy, Nicolas.....	205, 206
Risset, Jean-Claude.....	9, 45, 64, 65	Templeton, William B.....	221
Rogers, Carl.....	228	Throbbing Gristle.....	155
Rokeby, David.....	273-275, 277, 278	Tisseron, Serge.....	282
Roquigny, Anne.....	204	Tkaczyk, Viktoria.....	108
Rouffineau, Gilles.....	99, 175	Tokuyama, Tomonaga.....	107
Rousseau, Alain.....	86	Tolstoï, Léon.....	176
Rushkoff, Douglas.....	345	Turing, Alan.....	43, 227
Russolo, Luigi.....	9	Ussachevsky, Vladimir.....	66
Ruttmann, Walter.....	24, 26, 33	VanDerBeek, Stan.....	61, 76
Rybczynski, Zbigniew.....	143	Varèse, Edgar.....	9, 45
Sadeghi, Manoocheher.....	57	Vargoz, Frédérique.....	13, 14
Scanner.....	126, 179	Vasulka, Steina et Woody.....	72, 100, 101, 157, 158, 299, 323
Scemama, Céline.....	223	Vergo, Peter.....	23
Schaeffer, Pierre.....	9, 65, 128, 154, 230, 241, 300, 332	Vernhes, Eric.....	160-164
Schafer, Raymond Murray.....	140, 207, 208, 230	Vial, Stéphane.....	218
Schilingi, Jacopo Baboni.....	113, 114, 343	Vigo, Jean.....	128, 135
Schmitt, Antoine.....	89, 90, 92, 93, 304, 333, 342	Vigroux, Franck.....	92, 93, 333
Schönberg, Arnold.....	33	Viola, Bill.....	62, 72, 136, 268
Schubert-Minski, Theresa.....	111	Viola, Bill,.....	62, 72, 136, 268
Schubert, Alexander.....	111, 172, 250, 252-257, 342	Virgile.....	216
Schwartz, Lillian F.....	16, 20, 59, 61-67, 87, 297, 305	Walshe, Jennifer.....	254
Scriabine, Alexandre.....	22	Ward, Adrian.....	49, 89, 221, 229
Sedgwick, Edie.....	287-289, 291	Wardrip-Fruin, Noah.....	229
Semiconductor.....	16, 113, 115-118, 343	Warhol, Andy.....	287-289, 291
Shannon, Claude.....	227, 232	Weinbren, Grahame.....	170-172, 175, 176
Shepard, Roger.....	46, 64, 215, 291, 303		

Weizenbaum, Joseph.....	228
Wellek, Albert.....	337
Wertheimer, Max.....	302
Whitney, John et James. 15, 16, 19, 33-39, 42, 54-58, 64, 74, 75, 79-81, 83, 84, 86, 237, 297, 298, 300, 304, 316- 318, 341	
Wiene, Robert.....	161
Wilfred, Thomas.....	64, 87
Willoughby, Dominique.....	26, 27
Woolf, Virginia.....	140
Xenakis, Iannis.....	9, 46, 91, 331
Zajec, Edward.....	49
Zazeela, Marian.....	108
Zemanek, Heinz.....	111
Zénouda, Hervé.....	86, 87, 92
Zilio, Marion.....	246, 247, 258
Zuboff, Shoshana.....	244, 248, 249, 258
Zuckerberg, Mark.....	259

## Index rerum

- Abstraction.....11, 13, 15, 16, 19, 21, 40, 48, 64, 66, 114, 118, 121, 324, 339
- Accélééré.....62, 127, 135, 148, 164
- Accès 47, 54, 103, 116, 125, 146, 157, 166, 169, 193, 199, 208, 267, 341
- Acousmatique.....170, 174, 182, 286
- Acoustique.....126, 127, 157, 208, 231, 239, 270, 301
- ActionScript.....151, 152
- Agent conversationnel.....219, 229, 233-235
- Aléatoire....19, 43, 57, 69, 77, 82, 85, 122, 145, 146, 157, 180, 270, 276, 290, 316, 322, 323, 327, 335
- Algorithmique...8, 21, 46, 50-52, 61, 70, 78, 89, 91, 93, 132, 147-151, 165, 169, 177, 178, 207, 239, 248, 251, 255, 258, 310, 311, 347
- Analogie.....15, 23, 109, 313, 340
- Ancrage.....221, 310
- Animation CGI.....117
- Anthropomorphisation.....227, 229, 236, 292
- Appairage.....93, 109, 159, 188
- Archéologie des média.....10, 124, 126
- Art total.....22
- Art vidéo.....11, 17, 19, 20, 71, 186, 200, 217, 287, 339
- Artificialité.....226, 233, 240, 271
- Arts sonores.....10, 18
- Assistant vocal.....242, 244
- Audification.....117, 119
- Audio-game.....86
- Audiographique.....16, 17, 74, 82, 84, 104, 119, 296, 297, 302, 304, 307-310, 315, 318, 324, 326, 329, 330, 333, 335, 337, 341, 342, 344
- Audiovision.....8, 164, 208
- Audiovisual Environment Suite (AVES).....16
- Audiovisualité.....113, 144
- Audiovisuologie.....11
- Automate.....86, 227, 231
- Autonomie.....70, 89, 106, 263, 283, 285
- Autruiphanie.....218
- Axe-temps.....16, 124, 126, 134, 138, 145
- Back-spinning.....154
- Base de données....16, 102, 125, 151, 164-170, 178, 180, 181, 183-188, 190, 191, 193, 194, 199, 203, 204, 209, 240, 246, 255, 256, 261, 273, 282, 339, 341-343
- BASIC.....77
- Battement.....51, 110, 302, 343
- BEFLIX.....60-62, 66
- Biais 8, 47, 55, 67, 76, 101, 122, 219, 223, 240, 241, 246-248, 257, 342
- Binaire.....12, 13, 91, 102, 107, 187
- Blockchain.....345
- Boîte noire.....345
- Boucle.....75, 90, 106, 137, 138, 141, 154, 159, 195, 247, 270, 272, 274, 286, 287, 291, 341
- Bruit...8, 9, 15, 91, 97, 109, 110, 113, 116, 118, 124, 126, 134, 135, 137, 138, 140, 187-190, 195, 214, 220, 225, 240, 241, 272, 274, 290, 293, 308, 310, 311, 322-324, 334, 344
- Bruit blanc.....109, 322-324
- Bruitage.....15, 124
- Buchla.....71, 72
- Canon.....37, 143, 266
- Captteur.....21, 113, 138, 223, 249, 264, 265
- Capture.....16, 60, 75, 76, 85, 92, 96, 131, 135, 137, 162, 194, 202, 254-257, 260, 287-289, 296, 327, 329, 333, 341, 342
- Carte perforée.....43, 48
- Cartographie.....103, 203, 207-210, 218, 243, 249
- CD-Rom.....175
- CGU (Contenus Générés par les Utilisateurs)....149-151, 193, 221, 303
- Champ magnétique.....112
- Chatbot.....228, 237, 342
- Chiptune.....100
- Chœur virtuel.....264, 266
- Choralité.....17, 217, 218, 262, 267, 272, 284, 289
- Cinéma graphique.....15, 24, 26, 27
- Cinéma interactif.....186, 224
- Cinéma structurel.....147-149
- Clignotement.....51, 66, 109
- Cluster.....89, 197, 210

Coalescence.....	281	Écho	35, 140, 166, 197, 228, 266, 278, 310, 324, 330, 334, 344
Cocktail Party.....	294	Écologie.....	14, 17, 208, 213, 216
Codage.....	12, 99, 100, 130, 232, 256	Effet-ballet.....	29
Code	12, 13, 60-62, 77, 80, 89, 95, 97, 102, 105, 107, 138, 152, 168, 231, 246, 256, 291, 298, 304, 306, 307, 309, 317, 318, 333-335, 337, 342	Effet-clip.....	29
Code de création (Creative Coding).....	60, 80, 304	ELIZA.....	172, 228, 229, 263
Collage.....	341	Émotion.....	23, 55, 244, 247-249, 251, 253, 258, 261, 266, 269, 285
Combinatoire.....	18, 43, 68, 93, 212, 264	Enregistrement.....	16, 26, 34, 42, 116, 117, 125, 128, 130, 131, 134, 146, 170, 201, 207, 208, 212, 218, 226, 230, 242, 258, 272
Congruence.....	64, 67, 223	Esthétique programmée.....	67-69, 98
Consonance.....	302	Field Recording.....	203, 207, 208, 342
Continuum.....	188, 226, 293	FORTTRAN.....	48, 52, 56, 57, 76
Contrepoint	23, 33, 37, 58, 64, 74, 83, 182, 296, 310, 344	Fréquence	79, 87, 101, 109, 118, 126, 133, 134, 194, 196, 258, 270, 300, 302, 310, 314, 321, 322, 328, 344
Conversion.....	13, 103, 107, 110, 144, 340	Fugue.....	23
Correspondance.....	22, 23, 38, 53, 78, 79, 103-105, 123, 168, 184, 300, 306, 308, 310, 316, 317, 326, 328, 337, 339, 340	GAN (Generative Adversarial Network).....	256, 259
Couplage.....	109, 118, 132, 163, 203, 303, 306-310, 313	Générativité.	12, 17, 21, 47, 48, 50, 55, 57, 70, 71, 74-76, 78, 84, 85, 89, 91, 92, 94, 104, 113, 120-122, 132, 138, 142, 143, 153, 154, 161, 178, 181, 182, 185, 189, 227, 256, 259, 276, 278, 290, 297, 298, 305, 308, 318, 340
Cut-up.....	132	Gestalt.....	89, 91, 302, 303, 324
Datamining.....	258	Glissando.....	9, 35, 36, 324
Décrochage interactif.....	170, 173, 177	Glisser-déposer.....	149
Deep Fake.....	259, 269	Glitch.....	77, 124
Deep Learning.....	241, 254, 255	GPS.....	206, 209, 211, 343
Deep Time.....	115	GROOVE.....	63, 74
Détection	16, 110, 207, 236, 240, 245, 246, 248, 251-253, 258, 307, 342, 343	Hacking.....	99, 346
Détourage.....	143, 284, 342	Harmonie.....	54, 58, 80, 101, 111, 220, 330
Dialogue.	10, 12, 15, 66, 68, 69, 73, 93, 98, 128, 161, 177, 217, 218, 220, 226-229, 231-233, 244, 254, 274, 285, 287-289, 339, 340, 347	Harmonie digitale.....	54, 58
Digital Image Articulator.....	101	Harmonique.....	57, 302, 336
Digitalisation.....	13	Hoquet.....	282, 283
Direct Video Synthetizer.....	73, 100	HTML5.....	305
Discrétisation.....	13	Humanités numériques.....	12, 167, 242
Dissonance.....	302	Hyperdocument.....	167
Distorsion.....	156, 277, 287	Hypermédia.....	167-169, 199
Djing.....	125, 144, 154, 155, 161, 179, 186, 191, 218, 267, 292, 293, 342	Hypermusique.....	204
Doublage.....	221, 222, 257	Hypertexte.....	167, 168
Dramatisation.....	140	Illusion sonore.....	64, 291
Drone.....	87	Image opératoire.....	247
Échantillon	50, 90, 92, 126, 131, 155, 163, 242, 259, 260, 270, 272, 291	Image partagée.....	242
		Immatériel.....	216, 343

Improvisation.....	73, 92, 93, 149, 162-164, 255
Indexation....	102, 103, 168, 169, 180, 193, 241, 246, 249
Indifférence.....	13, 102, 105, 117
Inférence.....	261
Instrument...13, 25, 26, 31, 34, 43, 45, 46, 53, 70-72, 77, 81, 83-86, 91, 93, 101, 110, 120, 123, 131, 156, 162, 188, 203, 212, 222, 226, 253, 254, 273	
Intelligence artificielle. .52, 227-229, 235, 236, 241-244, 249, 250, 253	
Interactivité...17, 70, 74, 78, 85, 125, 137, 140, 143, 145, 148, 175, 185, 218, 227, 273, 341	
Interchangeabilité.....	175, 256, 336
Intervalle.....	36, 75, 122, 159, 298
Inversion.....	37, 125, 127-129, 145, 164, 223, 341, 344
Latent Space.....	256
Light Show.....	71
Linéarité.....	122, 166, 167, 341
Lingo.....	151, 152
Live A/V (Live Audiovisual Performance).....	185
Live Cinema.....	11, 17, 185, 342
Live Coding.....	96, 336
Logiciel.....16, 18, 21, 39, 45, 53, 61, 74, 78, 81-85, 103, 107, 117, 119, 120, 145, 148-154, 157, 158, 160, 163, 173, 183, 228, 232, 236, 237, 239, 243, 249, 254, 255, 265, 273, 282, 304, 323, 331-333, 337, 341, 345	
Lutherie.....	45, 87
Machine analytique.....	43
Mainframe.....	46, 52
Mappage (de données) .83, 103, 105, 112, 118, 123, 138, 310, 340	
Mapping.....	87, 103, 122, 138
Matérialité.....16, 126, 133, 213, 216, 217, 240, 344	
Metavers.....	345
Mickeymousing.....	29
Microphone....50, 130, 133, 156, 193, 203, 204, 265, 276	
Microstructure.....	125, 137, 145, 148, 188
MIDI.....	76, 77, 83, 103, 157
Mixage.....	92, 188, 203, 204, 210
Modélisation.....	17, 51, 118, 218, 223, 262
Modularité.....	106
Modulation de fréquence.....	87
Monitoring.....	343
Monologue.....	238, 239, 263, 287
Montage...15, 17, 24, 56, 57, 84, 103, 124, 128, 143, 146, 148-150, 152, 155, 159, 160, 163, 166, 169, 174, 175, 181-183, 185, 188, 189, 191, 194, 195, 202, 214, 221, 257, 270, 280, 281, 285, 287, 343	
Moog.....	63, 71
Morphing.....	253, 256, 277
Motion Bounce Effect.....	304, 324, 325
Mouvement cellulaire.....	50
Multimédia.....12, 151, 178, 186, 233, 236, 265, 272	
Music Mouse.....	78
Musique audiovisuelle.....	35, 79, 86, 297
Musique concrète.....	9, 50, 128
Musique électronique.....	9, 40, 66, 71, 75, 89, 186
Musique spectrale.....	9
Musique Visuelle (Visual Music)....11, 15, 18, 19, 21, 22, 27-29, 76, 79, 186, 296, 297, 319, 334, 339, 342	
Navigation...120, 125, 156, 169, 176, 190, 210, 212, 226, 270, 304, 326, 342, 343	
Net art.....	168, 345
NFT (non-fungible token).....	345
Non-chose.....	13
Notation.....	32, 85, 168
Numérisation.....	12, 246
Onde sinusoïdale.....	109
Opacité.....	242, 345
Opérateur.....77, 103, 131, 144, 145, 171, 172, 175, 185, 190, 205, 261	
Oramics.....	330, 331
Oscillateur.....	100, 306
Oscilloscope.....	40, 41
Panopticon.....	239
Parole....12, 124, 128, 166, 217-222, 224, 229, 230, 232, 233, 245, 266, 272, 278, 280, 283, 287-289, 292	
Partiel.....	194, 302, 310
Partition....22, 26, 29, 44, 52, 75, 86, 103, 128, 132, 159, 194, 212, 307, 322, 323, 331, 332	
Pellicule.....15, 30, 32, 100, 106, 127, 146, 160, 331, 339	
Performance audiovisuelle.....11, 12, 15, 88, 90, 94, 95, 250, 251, 342	
Performatio.....	144
Performativité.....	85, 125
PETSCII.....	96, 97

Phonographe. 69, 110, 126, 127, 134, 146, 201, 220, 230, 233	Scratching .....154, 161
Phonoscope.....219, 220	Selfie.....288, 341
Plotter.....15, 48	Signal.....11, 30, 66, 77, 92, 100, 101, 108, 110, 116, 130, 131, 133, 142, 156, 249, 270, 287, 322, 323, 346
Poésie sonore.....132	Silence..109, 127, 128, 136, 137, 139, 140, 217, 258, 276, 277, 289, 323
Polyphonie.....23, 194, 221, 272, 276, 283, 284	Snap.....17, 192-194, 196, 198, 342
Portrait ..17, 217-219, 221, 224-227, 237, 247, 259, 260, 263, 268, 284, 289-291, 341	Software.....13, 85, 148
Portrait parlant.....17, 219, 221	Solfège.....18, 21, 46, 298, 301
Précinéma.....15, 127, 146, 217, 219, 221, 342	Son delta.....128
Probabilité.....46, 203	Son optique.....32, 34, 36, 100, 128
Profilage.....258	Son organisé.....9
Programmabilité.....69, 70	Sonification.....77, 100, 103, 110-112, 117, 184, 341
Programmabilité.....70	Sonorisation.....29, 188, 310
Programmation. 16, 43, 51, 52, 56, 60, 61, 70, 76, 84, 90, 93, 95, 98, 99, 101, 120, 129, 139, 142, 151, 228, 277, 304-306, 316, 317, 320, 325, 331, 345	Sound Studies.....10
Prosodie.....240	Spatialisation. 74, 181, 211-213, 268, 289, 293, 306, 310, 320, 321, 337, 340, 342, 344
Raison computationnelle.....166, 167	Statua Citofonica.....239
Raison graphique.....166	Stéthoscope.....110, 113
Ralenti. 125-127, 134-141, 145, 146, 148, 158, 163, 164, 198, 289, 290, 293, 317	Stockage.....13, 43, 44, 129, 166, 273, 343, 344
Reconnaissance faciale.....246-248, 250, 251	Substance.....16, 66, 67, 86, 87, 135, 145, 148
Reconnaissance vocale.....198, 240, 258, 276, 277	SuperCollider.....121, 306
Réemploi (Found Footage).....16, 66, 76, 121, 125, 148, 150, 155, 161, 162, 164, 188, 190, 191, 194, 279, 282, 342	Supercut.....196
Requête.....9, 188, 282, 342	Surveillance...16, 126, 181, 199, 201, 207, 241, 244, 245, 247, 250, 251, 342, 343
Réseaux sociaux. 180, 192, 193, 198, 200, 201, 217, 288, 342	Synchronisation...8, 15, 37, 83, 104, 109, 124, 125, 149, 159, 188, 210, 219, 220, 222, 223, 233, 257, 263, 274, 290, 293, 294, 308, 309, 339, 341
Responsive.....8, 107, 305	Synthèse.....8, 9, 15, 18, 19, 26, 27, 30-36, 40, 41, 45, 46, 53, 64, 76, 79, 83, 84, 86-88, 100, 117, 118, 121, 143, 156, 183, 184, 186, 203, 227, 229, 232, 257, 263-267, 269, 270, 274, 301, 302, 305-308, 320, 322, 331, 340, 344
Rétrogradation.....128, 135, 136	Synthèse additive.....87, 88, 344
Rutt/Etra Scan Processor.....156	Synthèse granulaire.....87, 267, 269, 270
Rythme.....24-26, 33, 37, 66, 79, 95, 105, 109, 110, 115, 135, 141, 148, 154, 159, 183, 200, 220, 245, 258, 285, 287, 300, 302, 337, 344	Synthétiseurs vidéo.....71, 73
Saisie 16, 95, 103, 124-126, 134, 136, 143, 171, 210, 211, 215, 228, 341, 342	Tabularité.....167
Sampling.....160	Terres rares.....213
Saut (interactif)....29, 125, 145, 146, 148-154, 156, 157, 160, 163, 164, 173, 177, 270, 271, 291, 315, 341	Tête de lecture.....145-147, 151, 152, 154, 164, 190, 270, 331
Scam.....292-295, 342	Theremin.....31, 41, 71, 123
Scène auditive.....303	Timbre 23, 30, 32, 45, 53, 77, 79, 127, 217, 240, 258, 261, 300, 302, 309, 321, 328, 344

Tonalité.....	77, 82, 105, 111, 240, 302, 306, 318, 330, 331, 336, 337	Vectorisation.....	140
Traçabilité.....	258	Vibration.....	51, 126, 310, 334
Trace.....	48, 135, 191, 209, 272, 324, 329, 340, 344	Vidéodisque.....	170, 171, 226, 234
Traduction.....	76, 77, 100, 102-104, 112, 118, 246, 341	Visualisation	16, 42, 77, 85, 100, 103, 112, 114, 118, 181, 309, 327, 330, 341
Transduction.....	144, 163	Vjing.....	11, 89, 125, 161, 163, 186, 267
Transmutabilité.....	85	Vlog.....	17, 191, 198, 201, 242, 279-282, 342
UPIC.....	331	Voix off.....	144, 182, 201, 253, 285
Utilisateur....	70, 84, 85, 87, 108, 125, 145, 148-152, 168, 170, 178-180, 183, 192, 193, 196, 202, 207, 210, 212, 227, 228, 233, 236, 245, 318, 328, 331, 341, 345	Web audio.....	305
VAMPIRE.....	73, 74	Webcam.....	199-205, 223, 235, 279
Variabilité.....	70, 106, 127, 138, 144, 148, 181	Whitney-Reed RDTD.....	16, 81
		Zootrope.....	127
		Œuvre ouverte.....	68

## Table des illustrations

Fig. 1 : Paul Klee, Fuge in Rot (1921), © Media Art Net	21
Fig. 2 : Walter Ruttmann, <i>Opus 1</i> (1921), © See This Sound	24
Fig. 3 : Viking Eggeling, <i>Symphonie diagonale</i> (1921-1924), © RE:VOIR Vidéo	25
Fig. 4 : Hans Richter, <i>Rythme 21</i> (1921-1924), © RE:VOIR Vidéo	25
Fig. 5 : Oskar Fischinger, <i>Studie Nr 8</i> (1931), © Center for Visual Music	28
Fig. 6 : Oskar Fischinger, <i>Optical Poem</i> (1938), © Mubi	28
Fig. 7 : Oskar Fischinger, croquis de travail sur une partition, © Center for Visual Music	29
Fig. 8 : Piste optique sur film 35 mm, ©Wikipedia	30
Fig. 9 : Rudolf Pfenninger et ses motifs graphiques, @ Centre Pompidou	31
Fig. 10 : Oskar Fischinger et ses motifs graphiques, © Media Art Net	31
Fig. 11 : Les frères Whitney dans leur atelier , © Robert Russett, Cecile Starr (dir.), <i>Experimental Animation. Origins of a New Art</i> , Cambridge, Da Capo Press, 1976	34
Fig. 12 : À gauche le pendule, à droite les deux frères, © Robert Russett, Cecile Starr (dir.), <i>Experimental Animation. Origins of a New Art</i> , Cambridge, Da Capo Press, 1976	35
Fig. 13 : John et James Whitney, <i>Five Abstract Films Exercices</i> (1943-44), © The John Whitney Collection, DVD, Pyramid, 2009	36
Fig. 14 : Ben F. Laposky, <i>Oscillons</i> , circa 1953, © Centre Pompidou	40
Fig. 15 : Mary Ellen Bute (circa 1953), © Paul Young, <i>Le Cinéma expérimental</i> , Paris, Taschen, 2009	41
Fig. 16 : John Whitney, <i>Catalog</i> (1961), © Rhizome.org	42
Fig. 17 : IBM 704 Electronic Data Processing System (1957), © IBM, Computer History Museum	45
Fig. 18 : Frieder Nake, <i>105/130</i> (1965), © compArt, Center of excellence digital art	46
Fig. 19 : John Stehura, <i>Cybernetik 5.3</i> (1960-1965), © John Stehura, cyberanimation.tripod.com	50
Fig. 20 : John Stehura, <i>Cybernetik 5.3</i> (1960-1965), © John Stehura, cyberanimation.tripod.com	51
Fig. 21 : John Whitney, <i>Arabesque</i> (1975) , © The John Whitney Collection, DVD, Pyramid, 2009	54
Fig. 22 : Le Dr Jack Citron devant la console graphique IBM 2250, © The John Whitney Collection, DVD, Pyramid, 2009	54
Fig. 23 : Changement des valeurs de variables sur l'écran tactile grâce à un stylet, © The John Whitney Collection, DVD, Pyramid, 2009	54
Fig. 24 : John Whitney, <i>Experiments in motion graphics</i> (1968), © The John Whitney Collection, DVD, Pyramid, 2009	55
Fig. 25 : John Whitney, <i>Permutations</i> (1968), © The John Whitney Collection, DVD, Pyramid, 2009	56
Fig. 26 : Lillian F. Schwartz utilisant un joystick de contrôle, en 1974, © Center for Experimental Humanities	59
Fig. 27 : Captures d'un film 16 mm de documentation du BEFLIX, réalisé par Ken Knowlton en	60

1963. Ken Knowlton, *A Computer Technique for the Production of Animated Movies* (1963)

© HistoryofInformation.com

Fig. 28 : Lillian Schwartz, *Pixillation* (1970), avec l'assistance de Kenneth Knowlton, © Site web de Lillian Schwartz, lillian.com 63

Fig. 29 : Lillian F. Schwartz, *Googolplex* (1972), avec l'assistance de Kenneth Knowlton, © Site web de Lillian Schwartz, lillian.com 65

Fig. 30 : Stephen Beck et Warner Jepson, *Illuminated Music* (1973), © Ubu.com 72

Fig. 31 : Stephen Beck « jouant » de son Direct Video Synthesizer, © mmmmm.ch / Electronic Arts Intermix 72

Fig. 32 : Laurie Spiegel, © WFMU 73

Fig. 33 : Laurie Spiegel, *Voyages* (1977), réalisé avec VAMPIRE, © Laurie Spiegel YouTube 73

Fig. 34 : Malcolm Le Grice, *Arbitrary Logic* (1984-86), Capture d'écran, © REWIND Artists' Video 75

Fig. 35 : Malcolm Le Grice, *Arbitrary Logic* (1984-86), Quatre captures d'écran, © LUX 76

Fig. 36 : Laurie Spiegel, *Music Mouse* (1985-86), version pour Atari ST, © Laurie Spiegel, musicmouse.com 78

Fig. 37 : John Whitney, *Moon Drum* (1991), © Moon Drum, Twelve evocations by John Whitney, VHS, Mystic Fire Video, 1991 79

Fig. 38 : John Whitney, extraits de l'ouvrage *Digital Harmony. On the Complementarity of Music and Visual Art* (1980) 80

Fig. 39 : John Whitney utilisant le Whitney-Reed RDTD, avec des extraits de *Moon Drum* (1991), © *The John Whitney Collection*, DVD, Pyramid Media, 2009 81

Fig. 40 : Golan Levin, *Audiovisual Environment Suite* (1998-2000), © Golan Levin / flong.com 84

Fig. 41 : Golan Levin, *Audiovisual Environment Suite* (1998-2000), de gauche à droite : *Aurora* (1999), *Floo* (1999), *Yellowtail* (1999), *Warbo* (2000) et *Loom* (1999), © Golan Levin/The aesthetics + computation group 87

Fig. 42 : Golan Levin, *Scribble* (2000), performance audiovisuelle, © Golan Levin/The aesthetics + computation group 88

Fig. 43 : Antoine Schmitt, *Nanomachine* (2002-2004), performance audiovisuelle, © nouveauxmedias.net 90

Fig. 44 : Antoine Schmitt, *Nanoensembles* (2002), systèmes autonomes en ligne, © nouveauxmedias.net 90

Fig. 45 : Antoine Schmitt et Franck Vigroux, *Tempest @ MUTEK Montreal 2013*, © Hayeur 93

Fig. 46 : Antoine Schmitt et Franck Vigroux, *Tempest* (2012), © Galerie Charlot 93

Fig. 47 : Robert Henke, *CBM 8032 AV*, performance audiovisuelle (2016-2021), © Robert Henke 94

Fig. 48 : Robert Henke, *CBM 8032 AV* (2016-2021), début de la performance audiovisuelle, © Robert Henke 95

Fig. 49 : Robert Henke, *CBM 8032 AV* (2016-2021), captures de moments, © Robert Henke 96

Fig. 50 : Robert Henke, *CBM 8032 AV* (2016-2021), jeu de caractères PETSCII, © Robert Henke 97

Fig. 51 : Robert Henke, *CBM 8032 AV* (2016-2021), vocabulaire sonore. Planche de signaux utilisés, © Robert Henke 97

Fig. 52 : Cory Arcangel, <i>Super Mario Clouds</i> (2002), © Cory Arcangel	98
Fig. 53 : Cory Arcangel & Paper Rad, <i>Super Mario Movie</i> (2005), © Cory Arcangel / coryarcangel.com	98
Fig. 54 : Steina et Woody Vasulka, <i>The Matrix</i> , 1970-1972, © Fondation Langlois	100
Fig. 55 : Steina et Woody Vasulka, <i>Soundsize</i> , 1974, © MoMA	100
Fig. 56 : Ryoji Ikeda, <i>test pattern [n° 5]</i> , installation audiovisuelle (2013), © Ryoji Ikeda / ryojiikeda.com	104
Fig. 57 : Ryoji Ikeda, <i>test pattern [n° 4]</i> (2013), © Ryoji Ikeda / ryojiikeda.com	105
Fig. 58 : Ryoji Ikeda, <i>test pattern [100 m version]</i> (2013), © Ryoji Ikeda / ryojiikeda.com	105
Fig. 59 : Ryoji Ikeda, <i>test pattern [n° 1]</i> , installation (2008), © Ryoji Ikeda / ryojiikeda.com	107
Fig. 60 : SMPTE color bars, © Wikipedia	108
Fig. 61 : Philips PM5544 (SECAM variation), © Wikipedia	108
Fig. 62 : Différents modèles de stéthoscopes de Laënnec, © Wellcome Collection	110
Fig. 63 : Le Mailüfterl de Heinz Zemanek, © TU Wien Archive/Heinz Zemanek Estate	111
Fig. 64 : Tableau de données utilisé par Charles Dodge, © UMKC Special Collections	112
Fig. 65 : Jacopo Baboni Schilingi, <i>Argo</i> (2018), © Claude Mossessian/Grand Palais	113
Fig. 66 : Jacopo Baboni Schilingi, <i>Argo</i> (2018), une inspiration, en rouge, © Claude Mossessian/Grand Palais	114
Fig. 67 : Jacopo Baboni Schilingi, <i>Argo</i> (2018), une expiration, en bleu, © Claude Mossessian / Grand Palais	114
Fig. 68 : Jacopo Baboni Schilingi, <i>Argo</i> (2018), visualisation traditionnelle des données respiratoires récupérées, © Claude Mossessian / Grand Palais	114
Fig. 69 : Semiconductor, <i>Earthworks</i> (2016), © Semiconductor	115
Fig. 70 : Semiconductor, <i>Earthworks</i> (2016), données brutes, © Semiconductor	116
Fig. 71 : Semiconductor, <i>Earthworks</i> (2016), représentation graphique des données, © Semiconductor	116
Fig. 72 : Semiconductor, <i>Earthworks</i> (2016), exemple de géomodèle en couches, © Semiconductor	117
Fig. 73 : Semiconductor, <i>Earthworks</i> (2016), enregistrement d'un sismogramme, © Semiconductor	117
Fig. 74 : Semiconductor, <i>Earthworks</i> (2016), activité sismique organique (les magmas volcaniques), © Semiconductor	118
Fig. 75 : Semiconductor, <i>Earthworks</i> (2016), activité sismique glaciaire, © Semiconductor	118
Fig. 76 : Peter Beyls, <i>WindChime</i> (2006-2017), installation audiovisuelle en temps réel à partir de données en ligne, © Galerie Charlot	119
Fig. 77 : Peter Beyls, <i>WindChime</i> (2016-2017), Schéma de circulation des données, © Peter Beyls, « WindChime: database driven mapping Of Complex dynamical systems », 2014	122
Fig. 78 : Peter Beyls, <i>WindChime</i> (2016-2017), © Université catholique de Porto	123
Fig. 79 : Piotr Kowalski, <i>Time Machine II</i> (1981), © Piotr Kowalski / piotrkowalski.com	129

Fig. 80 : Piotr Kowalski, <i>Time Machine I</i> et installation galerie Feldman à New York (1979), © Centre Georges Pompidou/Musée national d'art moderne	130
Fig. 81 : Piotr Kowalski, Schéma du principe des <i>Time Machine</i> , © Centre Georges Pompidou/Musée national d'art moderne	131
Fig. 82 : Piotr Kowalski, <i>Time machine II (1981)</i> , © Piotr Kowalski / piotrkowalski.com	133
Fig. 83 : Thierry Kuntzel, <i>The Waves (2003)</i> , © Galerie de l'ésam, site de Caen	137
Fig. 84 : Thierry Kuntzel, <i>The Waves (2003)</i> , © Musée d'arts de Nantes	138
Fig. 85 : Thierry Kuntzel, <i>The Waves (2003)</i> , © Vidéoformes / Richard Beaune	138
Fig. 86 : Samuel Bianchini, <i>Temps libre (2004)</i> , © Samuel Bianchini / dispotheque.org	142
Fig. 87 : Samuel Bianchini, <i>Temps Libre, 2004</i> , © Samuel Bianchini / dispotheque.org	143
Fig. 88 : Nam June Paik, <i>Random Access (1963)</i> , © Guggenheim Museum	148
Fig. 89 : Barbara Lattanzi, <i>Optical De-dramatization Engine (O.D.E.) applied in 40-hour cycles to Thomas Ince's 'The Invaders', 1912 (2006)</i> , © Barbara Lattanzi / wildernesspuppets.net	148
Fig. 90 : Hollis Frampton, <i>Critical Mass (1971)</i> , © Anthology Film Archives	150
Fig. 91 : Ernie Gehr, <i>Serene Velocity (1970)</i> , © MoMA	150
Fig. 92 : Anne McGuire, <i>Strain Andromeda The (1992)</i> , © Mubi	150
Fig. 93 : Barbara Lattanzi, extrait du code source de <i>HC Critical Mass</i> © Barbara Lattanzi/wildernesspuppets.net	153
Fig. 94 : Barbara Lattanzi, <i>Optical De-dramatization Engine (O.D.E.) applied in 15-hour cycles to Ma-Xu Weibang's 'Ye ban gesheng' (Song at Midnight), 1937 (2008)</i> , © Barbara Lattanzi / Vimeo	154
Fig. 95 : Laurie Anderson et son Viophonograph, © Olivier Lussac / artperformance.over-blog.fr	156
Fig. 96 : Steina, <i>Violin Power (1969-78)</i> , © vasulkakitchen.org	158
Fig. 97 : Steina, <i>Violin Power The Performance (1992)</i> , © vasulka.org	159
Fig. 98 : Martin Arnold, <i>Passage à l'acte (1993)</i> , © Crac Occitanie	159
Fig. 98 : Martin Arnold, Partition de Pièce touchée, © Martin Arnold / Researchgate Max Neupert	160
Fig. 100 : Eric Vernhes, <i>L'interprète (2010)</i> , © Eric Vernhes / ericvernhes.com	161
Fig. 101 : Eric Vernhes, <i>L'interprète (2010)</i> , © Eric Vernhes / ericvernhes.com	163
Fig. 102 : Eric Vernhes, <i>L'interprète (2010)</i> , capture vidéo de différents moments d'improvisation, diffusés sur la colonne, © Vimeo / Eric Vernhes	163
Fig. 103 : Eric Vernhes, <i>L'interprète (2010)</i> , © Vimeo / Eric Vernhes	164
Fig. 104 : Grahame Weinbren et Roberta Friedman, <i>The Erl King (1982-85)</i> , © Vimeo / Grahame Weinbren	171
Fig. 105 : Grahame Weinbren et Roberta Friedman, <i>The Erl King (1982-85)</i> , © Alain Depocas, Jon Ippolito, Caitlin Jones (dir.), <i>L'approche des médias variables : la permanence par le changement</i> , New York / Québec, Guggenheim Museum/Fondation Daniel Langlois, 2003	173
Fig. 106 : Grahame Weinbren et Roberta Friedman, <i>The Erl King (1982-85)</i> , © Vimeo / Grahame Weinbren	176
Fig. 107 : Grahame Weinbren et Roberta Friedman, <i>Sonata (1993)</i> , © Vimeo / Grahame	177

## Weinbren

Fig. 108 : Lev Manovich, « Texas », <i>Soft Cinema</i> (2002-2005), © digitalartarchive.at	178
Fig. 109 : Lev Manovich, <i>Soft Cinema</i> (2002-2005), © digitalartarchive.at	181
Fig. 110 : Lev Manovich, <i>Soft Cinema</i> (2003), détail d'une visualisation de la base de données, qui présente des prises de vue géolocalisées à Berlin, © softcinema.net [archive.org]	182
Fig. 111 : Sergei Eisenstein, structure de montage d'une séquence de son film <i>Alexander Nevsky</i> (1939), © softcinema.net [archive.org]	182
Fig. 112 : Andreas Kratky, <i>Absences</i> (2004), © DVD <i>Soft Cinema</i> , MIT Press (2005)	184
Fig. 113 : Table de correspondances entre séquences vidéo et son, © Andreas Kratky, Livret du DVD <i>Soft Cinema</i> , MIT Press (2005)	185
Fig. 114 : SLIDERS_lab, <i>Sky Memory Project</i> , 2011-2014, © Jean-Marie Dallet / SLIDERS_lab	186
Fig. 115 : SLIDERS_lab, performance Résonances, IRCAM, Paris, France, 2006, © SLIDERS_lab	187
Fig. 116 : SLIDERS_lab, performance, IMAL Bruxelles, Belgique, 2008, © SLIDERS_lab	187
Fig. 117 : SLIDERS_lab, documentation du dispositif, © SLIDERS_lab	188
Fig. 118 : SLIDERS_lab, <i>TMWKTM</i> (2009-2015), © sliderslab.com	190
Fig. 119 : SLIDERS_lab, <i>Sky Memory Project</i> , 2011-2014, © Jean-Marie Dallet / SLIDERS_lab	191
Fig. 120 : Christian Marclay, <i>The Organ</i> (2019), © ARTnews / Benoît Florençon	192
Fig. 121 : Christian Marclay, <i>All Together</i> (installation de <i>Sound Stories</i> ), vue de l'installation au LACMA, Los Angeles, 2019, © Christian Marclay / LACMA	195
Fig. 122 : Christian Marclay, <i>Tinsel Loop</i> (installation de <i>Sound Stories</i> ), vue de l'installation au LACMA, Los Angeles, 2019, © Christian Marclay / LACMA.	196
Fig. 123 : Christian Marclay, <i>The Organ</i> (installation de <i>Sound Stories</i> ), vue de l'installation au LACMA, Los Angeles, 2019, © Christian Marclay / LACMA	197
Fig. 124 : Christian Marclay, <i>Talk to Me / Sing to Me</i> (installation de <i>Sound Stories</i> ), vue de l'installation au LACMA, Los Angeles, 2019, © Christian Marclay / LACMA	198
Fig. 125 : Christian Marclay, <i>Sound Tracks</i> (installation de <i>Sound Stories</i> ), vue de l'installation au LACMA, Los Angeles, 2019, © Christian Marclay / LACMA	199
Fig. 126 : Natalie Bookchin, <i>Location Insecure</i> (2005-2007), © Vimeo Natalie Bookchin	201
Fig. 127 : Jérôme Joy, <i>Nocinema</i> (1999-2004), © nocinema.org	203
Fig. 128 : Jérôme Joy, <i>Nocinema</i> (1999-2004), cartographie des sites de webcam, © nocinema.org	204
Fig. 129 : SLIDERS_Lab, <i>Sky Memory Net</i> (2015), © SLIDERS_Labs	205
Fig. 130 : Nicolas Tardy, <i>L'image contient peut-être</i> , © acousticcameras.org	206
Fig. 131 : Masaki Fujihata, <i>Field_work Alsace</i> (2002), © antiAtlas des frontières	210
Fig. 132 : Masaki Fujihata, <i>Simultaneous Echoes</i> (2009), © YouTube / Masaki Fujihata	212
Fig. 133 : Moritz Fehr, <i>Colosseum</i> (2015), images en 3D, © Moritz Fehr / moritzfehr.de	214
Fig. 134 : Moritz Fehr, <i>Colosseum</i> (2015), vues de l'installation, © Moritz Fehr / moritzfehr.de	215
Fig. 135 : Moritz Fehr, <i>Colosseum</i> (2015) © Moritz Fehr / moritzfehr.de	216
Fig. 136 : Georges Demenÿ, <i>Photographie de la parole</i> , © Gallica	220

Fig. 137 : Bruce Nauman, <i>Lip Sync</i> (1969), © Pinault Collection	223
Fig. 138 : Ingmar Bergman, <i>Monika</i> (1953), © Telerama	224
Fig. 139 : Luc Courchesne, <i>Portrait n° 1</i> (1990), © ZKM	225
Fig. 140 : Luc Courchesnes, <i>Portrait n° 1</i> (1990), © Fondation Langlois	226
Fig. 141 : Luc Courchesnes, <i>Portrait n° 1</i> (1990), © Media Art Net	226
Fig. 142 : Les Jeux des adolescents (les automates). Les têtes parlantes de l'abbé Mical., © Collections du Musée national de l'Éducation	232
Fig. 143 : Lynn Hershman Leeson, <i>DiNA</i> (2004-2006), © Lynn Hershman Leeson / lynnherhshman.com	235
Fig. 144 : Lynn Hershman Leeson, <i>Lorna</i> (1979-1983), © ZKM	236
Fig. 145 : Lynn Hershman Leeson, <i>CybeRoberta</i> (1996), © anthology.rhizome.org	236
Fig. 146 : Lynn Hershman Leeson,, <i>Electronic Diaries</i> (1986 —), © New Museum	236
Fig. 147 : Lynn Hershman Leeson, <i>Agent Ruby</i> (1998-2002), © SFMOMA	236
Fig. 148 : Différentes expressions d' <i>Agent Ruby</i> , © Fondation Daniel Langlois	236
Fig. 149 : Lynn Hershman Leeson, <i>DiNA</i> (2004-2006), © bitforms gallery	237
Fig. 150 : Lynn Hershman Leeson, <i>DiNA</i> (2004-2006), © SEQUITUR / Galerie Bridget Donahue	237
Fig. 151 : Lynn Hershman Leeson, <i>DiNA</i> (2004-2006), l'écran qui diffuse les « pensées » de <i>DiNA</i> , à trois moments différents, © Whitney Museum / YouTube	238
Fig. 152 : Kyriaki Goni, <i>Not allowed for algorithmic audiences</i> (2021), © Kyriaki Goni, copie privée	239
Fig. 153 : Kyriaki Goni, <i>Not allowed for algorithmic audiences</i> , 2021, © Kyriaki Goni, copie privée	240
Fig. 154 : Kate Crawford et Vladan Joler, <i>Anatomy of an AI System</i> (2018), détail., © anatomyof.ai	242
Fig. 155 : Jeff Guess, <i>Seamlessness</i> (2019), © MAK Center for Art and Architecture	244
Fig. 156 : Alphonse Bertillon, <i>Planche de visages de criminels</i> , © France Info	247
Fig. 157 : Trevor Paglen, <i>Machine-Readable Hito</i> (2017), © Trevor Paglen / paglen.studio	249
Fig. 158 : Trevor Paglen et le Kronos Quartet, <i>Sight Machine</i> , performance audiovisuelle (2018), © Cantor Center for Visual Arts at Stanford University	251
Fig. 159 : Trevor Paglen et le Kronos Quartet, <i>Sight Machine</i> , performance audiovisuelle (2018), © Barbican	252
Fig. 160 : Trevor Paglen, <i>Image Operations Op.10</i> (2018), © Trevor Paglen / paglen.studio	252
Fig. 161 : Alexander Schubert, <i>Convergence</i> (2020-21), version scénique, © Alexander Schubert / alexanderschubert.net	253
Fig. 162 : Alexander Schubert, <i>Convergence</i> (2020-21), version streaming, © Alexander Schubert / alexanderschubert.net	254
Fig. 163 : Capture de l'interface du logiciel d'apprentissage audio, développé avec l'Ircam pour <i>Convergence</i> , © Alexander Schubert / alexanderschubert.net	256
Fig. 164 : Heather Dewey-Hagborg et Chelsea Manning, <i>Probably Chelsea</i> (2017), © Benoît Montigné, Biennale Némo, octobre 2021	260
Fig. 165 : Bill Posters, <i>Big Dada/Public Faces</i> (2019), séquence avec Marina Abramović, ©	261

billposters.ch

Fig. 166 : Exemple de technologie Speech2face, © speech2face.github.io	261
Fig. 167 : Extrait documentation vidéo Nahas & Huitric (1993), © Archives vidéo de Paris 8	263
Fig. 168 : Monique Nahas et Hervé Huitric, <i>Masques et bergamasques</i> (1990), © Archives vidéo de Paris 8	264
Fig. 169 : Atelier de chœur virtuel, © Monique Nahas et Hervé Huitric / huitric-nahas.org	265
Fig. 170 : <i>Comme cela vous chante</i> (2002), Villette Numérique, © Monique Nahas et Hervé Huitric / huitric-nahas.org	266
Fig. 171 : Extrait documentation vidéo Nahas & Huitric (1993), © Archives vidéo de Paris 8	267
Fig. 172 : Granular Synthesis, <i>Modell 5</i> (1994-1996), © Ars Electronica	268
Fig. 173 : Granular Synthesis, <i>Modell 5</i> (1994-1996), GS-Retrospective, STRP Eindhoven, 2009, © Granular-Synthesis	270
Fig. 174 : Granular Synthesis, <i>Modell 5</i> (1994-1996), MAK_Museum of Applied Arts, Vienna, 1995, © Bruno Klomfar /Gebhard Sengmueller	271
Fig. 175 : Rafael Lozano-Hemmer, <i>Voice Array</i> (2011), © Rafael Lozano-Hemmer / lozano-hemmer.com	272
Fig. 176 : Rafael Lozano-Hemmer, <i>Voice Array</i> (2011), © Rafael Lozano-Hemmer / lozano-hemmer.com	273
Fig. 177 : Rafael Lozano-Hemmer, <i>Voice Array</i> (2011), © Rafael Lozano-Hemmer / lozano-hemmer.com	273
Fig. 178 : David Rokeby, <i>n-Cha(n)t</i> (2001), © David Rokeby / davidrokeby.com	274
Fig. 179 : David Rokeby, <i>n-Cha(n)t</i> (2001), Walter Phillips Gallery, Banff Centre for the Arts, © David Rokeby / www.davidrokeby.com	276
Fig. 180 : David Rokeby, <i>n-Cha(n)t</i> (2001), Walter Phillips Gallery, Banff Centre for the Arts, © David Rokeby / www.davidrokeby.com	276
Fig. 181 : Natalie Bookchin, <i>Testament</i> , 2009/2017, vue de l'installation multicanal, © Natalie Bookchin / bookchin.net	280
Fig. 182 : Natalie Bookchin, <i>My Meds</i> (2008), © Natalie Bookchin / bookchin.net	281
Fig. 183 : Natalie Bookchin, <i>Laid-Off</i> (2009), au moment de l'unisson, © Natalie Bookchin / bookchin.net	282
Fig. 184 : Candice Breitz, <i>Her</i> (1978-2008), © Goodman Gallery	285
Fig. 185 : Candice Breitz, <i>Him</i> , (1968-2008), © Goodman Gallery	285
Fig. 186 : Andy Warhol, <i>Outer and Inner Space</i> (1965), © The Andy Warhol Museum	289
Fig. 187 : Nicolas Clauss, <i>Terres arbitraires</i> (2014), © Nicolas Clauss / nicolasclauss.com	291
Fig. 188 : Nicolas Clauss, <i>Endless Portraits</i> (2015), © Nicolas Clauss / nicolasclauss.com	292
Fig. 189 : Joana Hadjithomas et Khalil Joreige, <i>The Rumour of the World</i> (2014), © Joana Hadjithomas et Khalil Joreige / hadjithomasjoreige.com	293
Fig. 190 : Joana Hadjithomas et Khalil Joreige, <i>The Rumour of the World</i> (2014), © Joana Hadjithomas et Khalil Joreige / hadjithomasjoreige.com	294
Fig. 191 : Capture de notre outil d'expérimentation en ligne, deux entités audiographiques évoluant en contrepoint sur deux parties d'écran, © audiographic lab	297

Fig. 192 : Manfred Mohr, <i>Matrix Elements</i> (1971), © ARC/Musée d'Art Moderne de la Ville de Paris	300
Fig. 193 : Armin Hofmann, <i>Manuel de design graphique</i> , 5 <sup>e</sup> édition, Niggli Verlag, 2004	302
Fig. 194 : Schéma de synthèse du processus de fabrication d'une animation audiographique	308
Fig. 195 : Exemple simple d'un code de génération d'une entité audiographique	310
Fig. 196 : Forme composite en mouvement, © audiographic lab	312
Fig. 197 : Courbe de Béziérs et bruit, © audiographic lab	312
Fig. 198 : « Rubans » tournants, © audiographic lab	312
Fig. 199 : Vassily Kandinsky, <i>Point et Ligne sur plan</i> (1926)	313
Fig. 200 : La surface d'un point statique, centré, croît, tandis que se joue un son continu, © audiographic lab	315
Fig. 201 : Un point se déplace de gauche à droite de l'écran, un son se joue et passe de l'enceinte gauche à droite, © audiographic lab	315
Fig. 202 : Dans l'étude ci-dessus, tous les points se déplacent de gauche à droite, sur une même trajectoire horizontale, à des vitesses aléatoires, © audiographic lab	317
Fig. 203 : John Whitney, <i>Experiments in Motion Graphics</i> (1968)	318
Fig. 204 : Étude audiographique d'après <i>Experiments in Motion Graphics</i> (1968), de John Whitney © audiographic lab	319
Fig. 205 : Oskar Fischinger, <i>Study n° 8</i> (1932), © Center for Visual Music	320
Fig. 206 : Oskar Fischinger, <i>Study n° 8</i> (1932), © Center for Visual Music	320
Fig. 207 : Reproduction programmée d'une scission, d'après un détail de <i>Study n° 8</i> (1932) d'Oskar Fischinger, © audiographic lab	321
Fig. 208 : Une fois la scission recrée sous forme d'objet informatique, nous pouvons la cloner, © audiographic lab	322
Fig. 209 : Dans cette étude on clique pour générer une forme sonore. Le son associé à la forme dépend de la position de la souris, © audiographic lab	322
Fig. 210 : Étude de bruit graphique et sonore, © audiographic lab	323
Fig. 211 : Différentes configurations de bruits graphiques et sonores, © audiographic lab	324
Fig. 212 : Reproduction de l'expérience du Motion Bounce Effect, © audiographic lab	325
Fig. 213 : Reproduction de l'expérience du <i>Motion Bounce Effect</i> , avec plusieurs paires de disques qui se croisent, © audiographic lab	326
Fig. 214 : Menu déployable pour naviguer dans les études audiographiques, © audiographic lab	327
Fig. 215 : Capture du défilement d'une page intégrant plusieurs études, © audiographic lab	328
Fig. 216 : Le mode « page », © audiographic lab	328
Fig. 217 : Le mode « hybride », © audiographic lab	328
Fig. 218 : Génération d'une « planche » qui correspond à la mise en correspondance du développement d'une forme complexe (composite) avec une gamme de fréquences., © audiographic lab	329
Fig. 219 : Façonnage d'une entité audiographique pour une étude de trace, © audiographic lab	330

Fig. 220 : Capture de l'interface de composition, en mode « timeline », © audiographic lab	330
Fig. 221 : Mode « édition », © audiographic lab	331
Fig. 222 : Mode « visualisation », © audiographic lab	331
Fig. 223 : Le système <i>Oramics</i> , de Daphné Oram (1957), © pointculture.be	331
Fig. 224 : Amit Pitaru, <i>Sonic Wire Sculptor</i> (2003), © YouTube d'Amit Pitaru	331
Fig. 225 : Le système <i>UPIC</i> , © ResearchGate Sylviane Sapir / Centre Iannis Xenakis	332
Fig. 226 : Le logiciel <i>Iannix</i> , ©iannix.org	332
Fig. 227 : <i>musigraphicode</i> , capture de l'interface de création et de composition, © audiographic lab	334
Fig. 228 : <i>musigraphicode</i> , Fête du code créatif, Centre Pompidou, 2018 © Benoît Montigné	335
Fig. 229 : La palette d'insertion d'animations prédéfinies, © audiographic lab	336
Fig. 230 : En haut la palette de <i>Live Coding</i> , en bas des exemples d'animation, © audiographic lab	337

## Extraits et documentations vidéo

Comme indiqué au début de la thèse, cette liste des extraits vidéo est également accessible sur un site compagnon : <http://quefaitlenumerique.benoit-montigne.net/>

Ils ont été tous consultés le 30/01/23.

### Première partie. Formes de la musique visuelle programmée

#### Chapitre 1. Abstractions graphiques et musicales du début du XXe siècle

Walter Rutmann, *Opus 1* (1921),

<http://www.medienkunstnetz.de/works/opus1/video/1/>, Media Art Net.

Viking Eggeling, *Symphonie Diagonale* (1921-1924), <https://vimeo.com/576389010>, Festival Punto y Raya.

Hans Richter, *Rythme 21* (1921-1924), <https://vimeo.com/262035970>, Festival Punto y Raya)

Oskar Fischinger, *Studie Nr 8* (1931), <https://vimeo.com/35735682>, Center for Visual Music.

Oskar Fischinger, *Optical Poem* (1938), <https://mubi.com/fr/films/an-optical-poem/trailer>, Mubi.

John Whitney, *Catalog* (1961), <https://rhizome.org/editorial/2008/nov/21/catalog-1961-john-whitney/>, Rhizome.org.

#### Chapitre 2. Le son des premiers films programmés

John Stehura, *Cibernetik 5.3* (1965), <https://www.facebook.com/watch/?v=2443493175964260>, Center for Visual Music.

John Whitney, *Permutations* (1968), [https://archive.org/details/permutations\\_201608](https://archive.org/details/permutations_201608), Internet Archive.

John Whitney, *Matrix I, II et III* (1971-1972), [https://archive.org/details/matrix\\_201703](https://archive.org/details/matrix_201703), Internet Archive.

John Whitney, *Arabesque* (1975), <https://digitalartarchive.siggraph.org/artwork/john-whitney-arabesque/>, SIGGRAPH.

Lillian F. Schwartz, *Pixillation* (1970), <http://lillian.com/1970-pixillation-4-min/>, site de Lillian F. Schwartz.

Lillian F. Schwartz, *Mutations* (1974), <http://lillian.com/films/>, site de Lillian F. Schwartz.

Lillian F. Schwartz, *Googolplex* (1972), <http://lillian.com/1972-googolplex-5-12-min/>, site de Lillian F. Schwartz.

### Chapitre 3. Inventer ses outils : des interfaces artistiques pour un art audiovisuel

Stephen Beck et Warner Jepson, *Illuminated Music* (1973), [https://ubu.com/film/beck\\_illuminated.html](https://ubu.com/film/beck_illuminated.html), Ubu.com.

Laurie Spiegel, *Voyages* (1977), <https://www.youtube.com/watch?v=0Fy5ilFIENo>, YouTube de Laurie Spiegel.

Malcolm Le Grice, *Arbitrary Logic* (1984-86), <https://www.youtube.com/watch?v=wCGIqRVRqmU>, LUX, à 5 min 48 s, sans le son.

John Whitney, *Moondrum* (1991), <https://archive.org/details/JohnWhitneyMoonDrum1991>, Internet Archive.

Golan Levin, *Audiovisual Environment Suite - AVES* (1998-2000), <https://acg.media.mit.edu/people/golan/aves/>, MIT.

Golan Levin, *Scribble* (2000), <http://flong.com/archive/projects/scribble/index.html>, site de Golan Levin.

Antoine Schmitt, *Nanomachine* (2002), <https://www.antoineschmitt.com/nanomachine-fr/>, site d'Antoine Schmitt.

Antoine Schmitt et Frank Vigroux, *Tempest* (2012), <https://www.youtube.com/watch?v=0z9zYpMTFnU>, Compagnie d'autres cordes.

Robert Henke, *CBM 8032 AV* (2016-2021), <https://www.youtube.com/watch?v=CY1RAs7x74w>, Compagnie d'autres cordes.

### Chapitre 4. Visualisation et sonification des données

Steina et Woody Vasulka, *The Matrix* (1970-1972), <https://www.fondation-langlois.org/html/f/page.php?NumPage=495>, Fondation Langlois.

Steina et Woody Vasulka, *Soundsizes* (1974), <https://www.fondation-langlois.org/html/f/media.php?NumObjet=12620&NumPage=465>, Fondation Langlois.

Ryoji Ikeda, *test pattern* (2008), version 100 m, <https://vimeo.com/120454782>, Vimeo de Ryoji Ikeda.

Jacopo Baboni Schilingi, *Argo* (2018), <https://vimeo.com/345013329>, Galerie Charlot.

Semiconductor, *Earthworks* (2016), <https://semiconductorfilms.com/art/earthworks/>, site de Semiconductor.

Peter Beyls, *WindChime* (2006-2017), <https://vimeo.com/128248373>, Présentation par Peter Beyls, avec extraits, Université catholique de Porto.

## Deuxième partie. Dramaturgies interactives

### Chapitre 1. Programmer l'axe-temps des images-son : inverser, ralentir, accélérer

Thierry Kuntzel, *The Waves* (2003), <https://www.youtube.com/watch?v=LL5czhM0Q>, YouTube La compagnie.

Samuel Bianchini, *Temps libre* (2004), <https://vimeo.com/75058234>, Vimeo de Samuel Bianchini.

### Chapitre 2. Épuisement du fragment audiovisuel : du saut interactif à la micro-structure

Barbara Lattanzi, *Critical Mass applied to Black Lives Matter* (2018), <https://vimeo.com/274382533>, Vimeo de Barbara Lattanzi.

Barbara Lattanzi, *Optical De-dramatization Engine (O.D.E.) applied in 40-hour cycles to Thomas Ince's 'The Invaders', 1912* (2006), <https://vimeo.com/idiomorphics>, Vimeo de Barbara Lattanzi.

Barbara Lattanzi, *Optical De-dramatization Engine (O.D.E.) applied in 15-hour cycles to Ma-Xu Weibang's 'Yeban gesheng' (Song at Midnight), 1937* (2008), <https://vimeo.com/141938927>, Vimeo de Barbara Lattanzi.

Barbara Lattanzi, *HF Critical Mass Soware HF Critical Mass soware applied to la revolución* (2015), <https://vimeo.com/137693423>, Vimeo de Barbara Lattanzi.

Steina, *Violin Power* (1969-78), <https://vimeo.com/555723409>, Vimeo The Kitchen.

Martin Arnold, *Passage à l'acte* (1993), <https://vimeo.com/398345542>, Vimeo L'Éclat.

Eric Vernhes, *L'interprète* (2010), <https://www.ericvernhes.com/project/linterprete/>, site d'Éric Vernhes.

### Chapitre 3. Esthétique des bases de données d'images et de sons

Grahame Weinbren et Roberta Friedman, *The Erl King* (1982-85), <https://vimeo.com/100314200>, Vimeo de Grahame Weinbren.

Grahame Weinbren, *Sonata* (1993), <https://vimeo.com/265094385>, Vimeo de Grahame Weinbren.

Lev Manovich, *Soft Cinema* (2002-2005), <https://vimeo.com/44602928>, V2\_.

SLIDERS\_lab, *Sliders\_1* (2005-2009), set 5, <https://vimeo.com/144611367>, et set 8, <https://vimeo.com/144611371>, Vimeo de SLIDERS\_lab.

SLIDERS\_lab, *TMWKTm* (2009-2015), [http://www.sliderslab.com/pages\\_fr/TRAVAUX/TMWKTm.html](http://www.sliderslab.com/pages_fr/TRAVAUX/TMWKTm.html), site de SLIDERS\_lab.

SLIDERS\_lab, *Sky Memory Project* (2011-2014), [http://www.sliderslab.com/pages\\_fr/TRAVAUX/Sky\\_Memory\\_Project.html](http://www.sliderslab.com/pages_fr/TRAVAUX/Sky_Memory_Project.html), site de SLIDERS\_lab.

Christian Marclay, *The Organ* (2019), <https://www.facebook.com/watch/?v=551036715723791>, Facebook du LACMA.

## Chapitre 4. Les images et les sons des paysages technologiques

Natalie Bookchin, *Location Insecure* (2005-2007), <https://vimeo.com/62458499>, Vimeo de Natalie Bookchin.

Natalie Bookchin, *Round the World* (2005-2007), <https://vimeo.com/62458499>, Vimeo de Natalie Bookchin.

SLIDERS\_Lab, *Sky Memory Net* (2015),  
[http://www.sliderslab.com/pages\\_fr/TRAVAUX/Sky\\_Memory\\_Net.html](http://www.sliderslab.com/pages_fr/TRAVAUX/Sky_Memory_Net.html), site de SLIDERS\_lab.

Nicolas Tardy, *L'image contient peut-être* (2021),  
<http://www.acousticcameras.org/playlist/nicolas-tardy/>, site Acoustic Cameras.

Masaki Fujihata, *Field\_work* Alsace (2002), <https://vimeo.com/86793927>, Vimeo de Caroline Bernard.

Masaki Fujihata, *Simultaneous Echoes* (2009), <https://www.youtube.com/watch?v=Wi32KBwDzMI>, Masaki Fujihata.

## Troisième partie.

### Dialogues et choralités. Les visages-voix de l'ère numérique contemporaine

#### Chapitre 1. Du portrait interactif à l'agent conversationnel

Bruce Nauman, *Lip Sync* (1969), <https://www.youtube.com/watch?v=hbT9GGJdKOo>, YouTube du GVA IVAM.

Luc Courchesne, *Portrait n° 1* (1990), version en ligne, <https://www.fondation-langlois.org/Portrait-no-1/>, Fondation Langlois et Capture vidéo,  
<https://vimeo.com/5827424>, Vimeo de Luc Courchesne.

Lynn Hershman Leeson, *DiNA* (2004-2006), documentaire vidéo, à 9 min 42 s,  
<https://www.youtube.com/watch?v=8Yggbr0cPXE>, bitforms gallery.

Kyriaki Goni, *Not allowed for algorithmic audiences* (2021),  
<https://vimeo.com/604759242>, Kyriaki Goni.

#### Chapitre 2. Visages et voix diagrammatiques

Jeff Guess, *Seamlessness* (2019), <https://www.guess.fr/works/seamlessness>, site de Jeff Guess.

Trevor Paglen & Kronos Quartet, *Sight Machine* (2018),  
<https://www.youtube.com/watch?v=5mg3MXETfj4>, YouTube Holland Festival.

Alexander Schubert, *Convergence* (2020-21),  
<https://www.alexanderschubert.net/works/Convergence.php>, site d'Alexander Schubert.

Jennifer Walshe et Memo Akten, *Ultrachunk* (2018), documentation vidéo,  
<https://vimeo.com/314016795n> Vimeo de Memo Akten, et captation de la performance, <https://www.memo.tv/works/ultrachunk/>, site de Memo Akten.

### Chapitre 3. Choralités mécaniques

Monique Nahas et Hervé Huitric, *Masques et bergamasques* (1990), <https://www.archives-video.univ-paris8.fr/video.php?recordID=305>, site des archives vidéo de Paris 8.

Monique Nahas et Hervé Huitric, documentations vidéo, 1993, animation des visages à l'aide de capteurs, <https://www.archives-video.univ-paris8.fr/video.php?recordID=315>, et digitalisation par technique Laser, <https://www.archives-video.univ-paris8.fr/video.php?recordID=314>, site des archives vidéo de Paris 8.

Granular Synthesis, *Modell 5* (1994-1996), <https://vimeo.com/43744967>, Vimeo d'Anti-Utopias.

Rafael Lozano-Hemmer, *Voice Array* (2011), <https://vimeo.com/34565650>, Vimeo de bitforms gallery.

David Rokeby, *n-Cha(n)t* (2001), <https://vimeo.com/17166940>, Vimeo de David Rokeby.

### Chapitre 4. Talking Heads

Natalie Bookchin, *Testament* (2009-2017), <https://vimeo.com/265980408>, et <https://vimeo.com/237790993>, Vimeo de Natalie Bookchin.

Candice Breitz, *Her* (1978-2008), <https://vimeo.com/72201223>, Vimeo de Candice Breitz.

Candice Breitz, *Him* (1978-2008), <https://vimeo.com/72198119>, Vimeo de Candice Breitz.

Andy Warhol, *Outer and Inner Space* (1965), <http://www.medienkunstnetz.de/works/outer-and-inner-space/video/1/>, Media Art Net, consulté le 30/01/23.

Nicolas Class, *Terres arbitraires* (2014), <https://vimeo.com/112859979>, Vimeo du bureau Olivia S. Sappey d'Anjou.

Joana Hadjithomas et Khalil Joreige, *The Rumour of the World* (2014), <https://www.youtube.com/watch?v=TTCJdfXk3ws>, YouTube du Jeu de Paume.

## Table des matières

<b>Sommaire.....</b>	<b>5</b>
<b>Introduction.....</b>	<b>8</b>
« Terme manquant : <b>sonore</b> ».....	8
Les relations entre image et son dans l'art.....	11
La difficile circonscription du numérique.....	12
Depuis le film sonore. Trois sillons d'exploration.....	15
Recherche-création : des études <i>audiographiques</i> .....	17
<b>Première partie.</b>	
<b>Formes de la musique visuelle programmée : de la composition d'abstractions visuelles et sonores à l'info-esthétique.....</b>	<b>19</b>
1.1 Abstractions graphiques et musicales du début du XXe siècle.....	21
1.1.1 De la musique visuelle.....	21
1.1.2 Naissance d'un cinéma graphique musical .....	24
1.1.3 Du graphique au sonore. La synthèse sonore optique ou la réversibilité d'un système.....	30
1.1.4 Un rééquilibrage de l'image et du son. La musique audio-visuelle des Frères Whitney.....	33
1.1.5 Sons et images des ondes. Physique et mathématiques de l'art.....	40
1.2 Le son des premiers films programmés.....	43
1.2.1 L'ordinateur avant l'ordinateur : une machine analytique .....	43
1.2.2 Émergence de la musique informatique et de l'art par ordinateur.....	44
1.2.3 <i>Cybernetik 5.3</i> , un système audio-graphique cellulaire.....	50
1.2.4 Harmonie digitale. Les films par ordinateur de John Whitney.....	54
1.2.5 Lillian F. Schwartz aux Bell Labs. Un mariage de l'image animée par ordinateur et de la musique informa- tique.....	59
1.2.6 Une esthétique programmée.....	67
1.3 Inventer ses outils : des interfaces artistiques pour un art audiovisuel.....	71
1.3.1 <i>Live Visuals</i> . Générer le visuel en temps réel.....	71
1.3.2 Malcolm Le Grice, <i>Arbitrary Logic</i> . Un orgue à couleurs génératif.....	75
1.3.3 L'aboutissement méconnu de la pensée de John Whitney : un outil de composition audio-graphique.....	79
1.3.4 Golan Levin. Une suite d'instruments audiovisuels.....	84
1.3.5 L'exploration de la relation audiovisuelle dans les formes semi-autonomes d'Antoine Schmitt.....	89
1.3.6 Générativité archéologique. Robert Henke et le CBM 8032.....	94
1.4 Visualisation et sonification des données.....	100
1.4.1 Traductions et transformations.....	100
1.4.2 Face aux données : l'indifférence de l'ordinateur.....	102

1.4.3 Ryoji Ikeda, <i>test pattern</i> . La perception des données par l'hypersynchronisation audiographique.....	104
1.4.4 La sonification. Des sciences à l'art.....	110
1.4.5 Audiographie du corps humain.....	113
1.4.6 Géologie vidéo-sonique : <i>Earthworks</i> , de Semiconductor.....	115
1.4.7 Peter Beyls. Des systèmes audiographiques auto-organisés.....	119

## Deuxième partie.

### Dramaturgies interactives, dialectique des images et des sons dans les média

#### numériques..... 124

2.1 Programmer l'axe-temps des images-son : inverser, ralentir, accélérer.....	126
2.1.1 Archéologie des sons et des images animées inversés.....	126
2.1.2 La <i>Time Machine</i> de Piotr Kowalski. Renversements temporels de l'image et du son.....	129
2.1.3 Pratiques fondatrices du ralenti.....	134
2.1.4 <i>The Waves</i> , de Thierry Kuntzel. Passages du mouvant à l'immobile, du sonore au silence.....	136
2.1.5 <i>Temps Libre</i> , de Samuel Bianchini. La voix agitatrice.....	141
2.2 Épuisement du fragment audiovisuel : du saut interactif à la microstructure.....	145
2.2.1 La figure du « saut » dans les arts interactifs.....	145
2.2.2 Barbara Lattanzi. Un cinéma structurel algorithmique .....	147
2.2.3 Le <i>scratching</i> ou l'art de la tête de lecture.....	154
2.2.4 <i>L'interprète</i> , d'Eric Vernhes. Vers le sampling audio-vidéo.....	160
2.3 Esthétique des bases de données d'images et de sons.....	164
2.3.1 La base de données dans l'art.....	165
2.3.2 Une figure fondamentale. Le décrochage interactif et acousmatique dans <i>The Erl King</i> .....	170
2.3.3 <i>Soft Cinema</i> . Le cinéma algorithmique de Lev Manovich.....	177
2.3.4 Dé-monter et remonter le cinéma. Les performances audiovisuelles de SLIDERS_lab.....	185
2.3.5 Le son des <i>Vlogs</i> et l'analyse automatique. <i>Sound Stories</i> , de Christian Marclay.....	191
2.4 Les images et les sons des paysages technologiques.....	199
2.4.1 Images et sons du monde en direct : du voyage imaginaire à la surveillance.....	199
2.4.2 Paysages visio-sonores : des caméras et des micros sur le monde.....	202
2.4.3 Fragments géolocalisés. Les <i>Field_works</i> de Masaki Fujihata.....	209
2.4.4 <i>Colosseum</i> . Matérialité du numérique et écologie des média.....	213

## Troisième partie.

### Dialogues et choralités. Les visages-voix de l'ère numérique contemporaine.....217

3.1 Du portrait interactif à l'agent conversationnel.....	219
3.1.1 Figures du portrait parlant.....	219
3.1.2 C'est moi que vous regardez ? Un portrait interactif.....	224
3.1.3 Les <i>chatterbots</i> . Premiers dialogues avec l'ordinateur.....	227
3.1.4 Détacher la voix du corps : paroles artificielles.....	229

3.1.5 <i>DiNA</i> , un agent conversationnel de Lynn Hershman Leeson.....	234
3.1.6 Sur écoute : confessions d'un assistant artificiel.....	238
<b>3.2 Visages et voix diagrammatiques.....</b>	<b>241</b>
3.2.1 Anatomies de la voix comme commande.....	241
3.2.2 Détections biométriques de la déviance : typage des visages et reconnaissance faciale.....	246
3.2.3 La détection sonore et visuelle comme dispositif de composition.....	250
3.2.4 Du visage à la voix, de la voix au visage : clones et chimères numériques.....	257
<b>3.3 Choralités machiniques.....</b>	<b>262</b>
3.3.1 Chœurs virtuels. Opéra d'ordinateurs.....	262
3.3.2 <i>Modell 5</i> , de Granular Synthesis. Grains de visage-voix.....	267
3.3.3 Confessions à des lumières de voix.....	271
3.3.4 Machines en(chantées) vocales. <i>n-Cha(n)t</i> , de David Rokeby.....	273
<b>3.4 Talking Heads.....</b>	<b>279</b>
3.4.1 Schèmes de l'extimité. Autour du travail de Natalie Bookchin.....	279
3.4.2 Candice Breitz. Théâtres noirs, une esthétique de l'isolation.....	284
3.4.3 Silences du regard.....	289
3.4.4 <i>The Rumour of the World</i> . Une anthropomorphisation des scams.....	292

## Quatrième partie.

### **Carnet de recherche-création. Une plate-forme expérimentale web pour la recherche audiographique ..... 296**

4.1 Des études pour l'exploration de l'art « audiographique ».....	297
4.1.1 Un art « audiographique ».....	297
4.1.2 Décloisonner les fondamentaux de l'animation et du sonore.....	297
4.1.3 Le noir et blanc.....	299
4.1.4 Les manuels d'études.....	300
4.1.5 Psychologie de la forme : vers une Gestalt <i>audiographique</i> .....	302
4.2 Un objet informatique audiographique : approche conceptuelle et technique .....	304
4.2.1 La réalisation d'études en ligne.....	304
4.2.2 Langages et techniques : HTML5, JavaScript et Web Audio.....	305
4.2.3 Des moteurs de synthèse générative .....	305
4.2.4 Configuration de l'entité audiographique.....	308
4.3 Quelques exemples d'études.....	310
4.3.1 Une étude fondamentale : « un point animé, un son ».....	312
4.3.2 Isolation et reprogrammation de micro-séquences : exemples avec Oskar Fischinger et John Whitney...316	
4.3.3 Notions transversales. L'exemple d'une étude de bruit.....	322
4.3.4 Approche perceptive : autour d'une illusion <i>audiographique</i> .....	324

4.4 Des interfaces : façonnage, navigation et composition.....	326
4.4.1 Navigation dans une page d'études.....	326
4.4.2 Les trois modes de consultation d'une étude.....	327
4.4.3 Espaces de façonnage.....	328
4.4.4 Une <i>timeline</i> de composition.....	329
4.4.5 Expérimentations sur les interfaces.....	330
4.5 Une version dédiée au jeune public : musigraphicode.....	333
4.5.1 Un atelier d'art <i>audiographique</i> et des projections.....	333
4.5.2 Fonctionnalités de <i>musigraphicode</i> .....	335
4.5.3 Tonalité, rythmique, interchangeabilité.....	336
4.5.4 Des correspondances « primitives ».....	337
<b>Conclusion.....</b>	<b>339</b>
L'espèce « cinéma sonore » à l'ère du numérique.....	339
Des correspondances à l'apprentissage automatique.....	340
Figures émergentes de l'image-son.....	342
Ouverture de notre recherche-crédation.....	344
Face aux interfaces invisibles. Quelles stratégies artistiques ?.....	345
Images et sons après le numérique : mutation, glissement, rupture.....	346
<b>Bibliographie.....</b>	<b>348</b>
<b>Sitographie.....</b>	<b>356</b>
<b>Index nominum.....</b>	<b>363</b>
<b>Index rerum.....</b>	<b>368</b>
<b>Table des illustrations.....</b>	<b>373</b>
<b>Extraits et documentations vidéo.....</b>	<b>382</b>
<b>Table des matières.....</b>	<b>387</b>
<b>Annexes.....</b>	<b>391</b>
Portfolio. Activités durant la période de thèse.....	391
Entretiens et articles publiés.....	395

## Annexes

### ***Portfolio. Activités durant la période de thèse***

#### **Enseignement et recherche**

##### **Enseignant Ensad en design interactif et code de création (depuis 2005)**

Licence Design graphique, École nationale supérieure des Arts Décoratifs de Paris

« L'objectif du cours est la conception et la réalisation de projets multimédia. Sur la base d'un apprentissage des langages fondamentaux du Web et grâce aux nouveaux outils de prototypage, l'étudiant s'initie aux différents aspects du design interactif : design d'interfaces, expérience utilisateur, responsive design, animation graphique, création sonore. En parallèle est développée une approche de la programmation à travers la pratique du code de création (expérimentations génératives interactives). Sont explorés objets éditoriaux, médias interactifs, sites institutionnels, applications en ligne, navigation dans des corpus, cartographies interactives, œuvres en ligne, nouvelles narrations... Autour des sujets principaux s'articulent des moments consacrés à la culture du design interactif et des arts numériques (histoire, actualité, usages et tendances). Ils donneront lieu à des ouvertures théoriques et critiques sur l'emploi des nouvelles technologies dans les hypermédias artistiques et les champs multiples du design graphique contemporain. »

##### **Membre du groupe de recherche « Formes du mouvement » (2021-2023)**

Ensadlab/École nationale supérieure des Arts Décoratifs de Paris

Coordination Serge Verny

« Comprendre et manipuler le mouvement dans ses dimensions physiques, perceptives ou imaginaires sont des enjeux qui ont intéressé tant les scientifiques que les artistes. Ainsi, en fonction des nécessités, des besoins ou des aspirations, différents savoirs, ou savoir-faire se sont dégagés dans des efforts de description, de compréhension, d'appropriation ou de transformation. Chaque discipline se dotant d'un langage et de méthodes pour explorer ce territoire, en constitue ainsi un éclairage particulier. Face à cette diversité, l'étudiant, voire le professionnel, peut se sentir dépassé. Aussi, l'enjeu de ce projet de recherche est double. Est-il possible de proposer un outil, d'abord à vertu pédagogique, qui permette d'avoir une vision d'ensemble de ces différentes approches, de se repérer dans ce territoire et d'en comprendre les liens et les passages ? Mais aussi quelles nouvelles régions, quelles frontières inexplorées apparaîtront alors ? »

##### **Chargé de cours Arts audiovisuels (2017-2019)**

Licence 2 Arts plastiques, Université de Paris 8

« Depuis la musique visuelle jusqu'aux pratiques contemporaines de la performance audiovisuelle, en passant par l'art vidéo, des artistes développent des langages esthétiques dans lesquelles les composantes sonore et visuelle dialoguent et interagissent sans que l'une n'asservisse l'autre. Nous explorerons les nouvelles modalités de ces formes d'art hybrides à la lumière du travail des artistes pionniers ».

Œuvres de : *Oskar Fischinger, Len Lye, John et James, Whitney, Woody & Steina Vasulka, Lillian Schwartz, Peter Kubelka, Norman McLaren, Michael Snow, Jean-Luc Godard, Christian Marclay, Ryoji Ikeda, Robert Henke...*

Parmi les champs couverts : musique visuelle, animation expérimentale, cinéma expérimental, danse, art vidéo, arts numériques, arts interactifs, écrans, scène, projections, lumières, interactivité, numérique, temps réel, performance audiovisuelle...

## Chargé de cours magistral en arts sonores (2015 à 2017)

Licence 1 Arts & et médias, Université Paris 3 — Sorbonne Nouvelle

« Si l'art sonore est un genre, il conviendra d'identifier la place singulière qu'il occupe, dans un entre-deux permanent, explorant les frontières des disciplines (musique, arts plastiques, arts numériques, poésie, art radiophonique, danse, cinéma, théâtre, architecture), cultivant une ambivalence subversive typique de la période contemporaine. »

Œuvres de : *Max Neuhaus, Alvin Lucier, Rolf Julius, Bernhard Leitner, Laurie Anderson, Bernard Heidsieck, Harry Bertoia, Len Lye, Céleste Boursier-Mougenot, Bill Fontana, La Monte Young...*

Parmi les champs couverts : sculpture sonore, installations sonores, son dans l'art contemporain, arts audiovisuels, poésie sonore, musiques expérimentales...

## Bourses

### Obtention d'une bourse FSIR pour *audiographic lab* (janvier 2017)

« Le Fonds de soutien à l'initiative et à la recherche (Fsir) a pour objectif principal le soutien des initiatives d'expérimentation et de recherche artistique et/ou culturelle, qui participent au renouvellement des formes et des esthétiques, en favorisant les projets menés en coopération. Les projets soutenus devront concerner les domaines suivants : recherche et développement, mutualisation et/ou coopération. »

## Présentations/Expositions

### Présentation de notre application graphique et sonore, *musigraphicode*, et atelier jeune public pour la Fête du Code Créatif au Centre Pompidou (4 et 5 mars 2018)

« Notre application web *musigraphicode* donne à l'enfant différentes possibilités de générer des formes graphiques abstraites et des motifs sonores puis de faire varier leurs paramètres en direct. Cet outil est une version dérivée d'*Audiographic Lab*, notre plateforme en ligne dédiée à l'exploration de la création audiographique. »

## Organisation d'événements

### Coordination du colloque « Architectures de mémoire » (novembre 2015)

Musée d'art et d'histoire de Saint-Denis/Bibliothèque de l'Université de Saint-Denis/Archives nationales de Pierrefitte-sur-Seine, organisé par Jean-Marie Dallet (Paris 8) et Bertrand Gervais (UQAM), dans le cadre du projet « Architectures de mémoire et multidiffusion » du Labex Arts-H2H.

« Le modelage des bases de données et leur exploration constituent, suivant les mots de Lev Manovich, une nouvelle catégorie-clé de la culture, l'«info-esthétique». Cette esthétique de l'information est fondée sur la description sémantique des contenus, sur leur lisibilité en somme, dont le moteur de recherche serait l'outil privilégié. La mise en forme de l'information, langagière mais aussi spatiale, apparaît alors comme une nouvelle façon de décrire et de représenter le monde. »

**Coordination de la journée d'étude « Architectures de mémoire » (novembre-décembre 2014)**

Université de Paris 8, Saint-Denis, organisée par Jean-Marie Dallet (Paris 8) et Bertrand Gervais (UQAM), dans le cadre du projet « Architectures de mémoire et multidiffusion » du Labex Arts-H2H.

**Communications****Interview par David Christoffel (mars 2018)**

Interview par David Christoffel, dans l'émission Magnétique d'Espace 2, autour de la thématique des logiciels d'apprentissage en musique.

**Intervention. « audiographic lab, une plateforme en ligne dédiée à la découverte de la création audiographique » (15 octobre 2016)**

Le Mabilay, Rennes, dans le cadre de Demain #2, les rencontres de la créativité numérique, festival Maintenant, musiques, arts et technologies.

« Demain ! est un espace de réflexions et d'échanges autour de la créativité numérique. à travers conférences, tables rondes, workshops et découvertes de projets artistiques innovants, sont mises en lumière les initiatives des acteurs et professionnels d'aujourd'hui qui font et participent à la création de demain. »

**Intervention. « audiographic lab, une plate-forme web adaptative pour l'étude des interactions entre création sonore et image graphique en mouvement » (3 et 4 décembre 2015)**

Maison de la Recherche de l'Université Paris 3, dans le cadre du colloque International « Sons mobiles » conçu et organisé par Laurence Allard, Roger Odin, Laurent Creton IRCAV-Paris 3.

« Les 3 et 4 décembre 2015 pour la quatrième saison du colloque annuel de "Mobile et Création" (IRCAV-Paris 3/Labex ICCA) ambitionne de rassembler des chercheurs et des professionnels autour de la dimension sonore des usages créatifs des terminaux mobiles (des téléphones simples aux appareillages les plus sophistiqués : smartphones, tablettes, phablettes et autres engins mobiles). »

**Conférence en images. « Musique live/Images live, histoires de rencontres », avec Gilles Alvarez (10 octobre 2015)**

Philharmonie de Paris, Paris, dans le cadre de Nêmo, biennale internationale des arts numériques.

**Intervention. « Représentation audiographique expérimentale » (26 septembre 2014)**

Ircam, Paris, dans le cadre de la journée d'études « la visualisation des données numériques », organisée par Samuel Goldszmidt et Rémy Campos dans le cadre du consor-

tium Musica (Huma-Num) dans le cadre de la 2e rencontre sur les bases de données et les ressources numériques.

## **Animation**

### **Fondateur et animateur du site *Sonore Visuel*, dédié à l'actualité des arts sonores et audiovisuels (depuis 2012)**

« Sonore Visuel est consacré à l'histoire et à l'actualité des arts audiovisuels dans leur acception la plus large : arts sonores, performance audiovisuelle, installation sonore, sculpture sonore, art vidéo, musique visuelle, cinéma expérimental, animation, art vidéo, nouveaux médias, transmédia. Le site explore le rapport image/son dans l'art à travers une sélection d'artistes et d'œuvres. »

## **Formations**

### **Formation. Design d'expériences audio sur le Web (13 au 15 mars 2018)**

Ircam, « Concevoir et réaliser des performances et des installations participatives s'appuyant sur les technologies du web ».

## **Autres activités**

### **Membre du réseau thématique pluridisciplinaire SON:S lancé par l'Institut des sciences humaines et sociales du CNRS, coordonné par Karine Le Bail (depuis 2019)**

« SON:S, réseau thématique pluridisciplinaire (RTP) lancé en 2019 par l'Institut des sciences humaines et sociales du CNRS (InSHS), connecte de manière inédite les équipes et laboratoires ayant le son et l'écoute pour principal territoire de recherche. Provoquant des situations d'échanges et d'interactions originales et encourageant la construction de programmes de travail collectifs, SON:S vise à faire émerger une communauté d'attention. »

### **Membre du Jury pour la thèse SACRe/PSL/ENS d'Hadrien Jean (décembre 2018)**

Soutenance de la thèse de doctorat « Apprentissage perceptif dans le domaine de la hauteur et de l'attention sélective auditive », programme SACRe de l'université PSL (Paris Sciences et Lettres), préparé à l'École Normale Supérieure (ENS), sous la direction de Daniel Pressnitzer.

## ***Entretiens et articles publiés***

Les deux articles listés ci-dessous sont disponibles sur les pages suivantes.

### **Entretien. « L’art audiographique : la “musique visuelle” à l’ère du numérique » (2020)**

Revue Culture et Recherche n° 141, « Cinéma, audiovisuel, son », printemps-été 2020, 96 pages. Dossier coordonné par : Jean Carrive (INA), Géraldine Poels (INA), Barbara Turquier (La Fémis), avec la collaboration de Solène Bellanger (Ministère de la Culture/DGCA).

« Ce numéro propose un panorama des recherches les plus récentes menées au sein des établissements du ministère de la Culture dans les domaines du cinéma, de l’audiovisuel, des arts visuels et multimédias, de la musique et des œuvres sonores. Il témoigne de la spécificité des pratiques de recherche dans ces établissements, où elles s’articulent avec la création et la formation, l’archive et le patrimoine, ainsi qu’avec les mondes professionnels de la culture.

Les articles, rassemblés selon quatre grandes thématiques, reflètent la diversité des programmes de recherche sur les images animées et les sons : approches transdisciplinaires mobilisant les outils liés aux humanités numériques pour renouveler l’analyse de vastes corpus ; travaux sur la place des images et des sons dans la société, sur les usages des produits médiatiques, leurs enjeux politiques et économiques ; projets de recherche-crédation développés en particulier dans les écoles d’art et de cinéma ; enfin recherches à des fins de conservation et de restauration d’œuvres cinématographiques, audiovisuelles, musicales, ou encore à des fins d’archivage du Web. »

### **Article. « Chercher l’équilibre : dialogues contemporains entre image et son » (juin 2015)**

Article pour la revue annuelle de saison de l’Ensemble intercontemporain, dans le cadre du dossier « Les nouvelles scènes de la musique ».

# L'art audiographique : la « musique visuelle » à l'ère du numérique

**Entretien avec Benoît Montigné**, enseignant en multimédia à l'École nationale supérieure des arts décoratifs, membre du groupe de recherche Formes du mouvement de l'Ensadlab, chargé de cours en arts audiovisuels et chercheur-doctorant TEAMeD / AI-AC à l'université Paris 8.

*Culture et Recherche : Vous poursuivez actuellement un doctorat en esthétique, sciences et technologie des arts au sein du laboratoire AI-AC de l'université Paris 8<sup>1</sup> et en collaboration avec l'Ensadlab<sup>2</sup>. Votre recherche-crédation est axée sur l'étude des relations interactivité / image / son, ce que vous appelez l'art audiographique. De quoi s'agit-il ?*

**Benoît Montigné :** Depuis les origines du cinéma se développent des formes d'art hybrides mêlant formes abstraites en mouvement et son. Dans des écrits datant de 1919 et publiés de manière posthume, Walter Ruttmann imagine une nouvelle forme d'art, qui se situe « aux franges de la peinture et de la musique<sup>3</sup> » ; le développement temporel en constitue l'essence formelle. Il met en pratique ses théories deux années plus tard et devient l'un des pionniers de ce courant émergent avec son film *Lichtspiel Op. 1* (1921). Souvent apparentées à ce que l'on a appelé la « musique visuelle », les œuvres dans lesquelles sont mis en scène des contrepoints de formes abstraites et de formes sonores n'ont jamais cessé de se développer, en témoignant les courts-métrages des

frères Whitney dans les années 1940, les travaux vidéo de Steina et Woody Vasulka dans les années 1960, ou plus récemment les installations numériques de Ryoji Ikeda. Aujourd'hui, grâce aux possibilités croissantes de générativité et d'interactivité en temps réel et par le biais des nouveaux dispositifs de spatialisation sonore – tel le son 3D – et de ceux dédiés à la projection – comme le *mapping vidéo* –, il est raisonnable de penser que ces formes continueront d'évoluer.

Différents langages coexistent dans ces œuvres hybrides, aussi me semblait-il pertinent, dans une perspective d'analyse, d'adopter une approche transversale : repartir des fondements de chacun de ces champs et les croiser pour en dégager une grammaire plastique. J'ai ainsi élaboré un outil de recherche-crédation pour explorer ce que je nomme l'art *audiographique*<sup>4</sup>. Il se présente sous la forme d'une plateforme Web ([www.audiographiclab.com](http://www.audiographiclab.com)) qui met à profit les technologies Web audio récentes du HTML5, ce qui le rend très accessible.

*C&R : Comment faites-vous interagir formes graphiques et formes sonores ?*

**B.M. :** Les recueils d'« études » (au sens de pièces d'apprentissages) sont légion dans la musique, mais il en existe également dans les arts graphiques. L'un des plus intéressants reste pour moi le *Manuel de création graphique*<sup>5</sup> du designer Armin Hofmann, qui propose une série d'études abstraites, toutes en noir et blanc, basées sur des compositions de formes graphiques basiques comme le point ou la ligne. L'ensemble constitue une forme d'abrégé de son enseignement à l'école de design de Bâle. Il y explore différentes polarités graphiques liées à la trame, la densité, ou bien encore la distribution.

Sur ce modèle, j'ai imaginé une série d'études interactives à l'écran dans lesquelles se croisent les fondamentaux de la création sonore (variations d'intensité, de fréquence, de composition spectrale, jeux de

1. Équipe TEAMeD « Théorie Expérimentation Arts Médias et Design », laboratoire de recherche AI-AC, « Arts des images et art contemporain », sous la direction de Jean-Marie Dallet.

2. Groupe de recherche « Formes du mouvement », dirigé par Nicolas Nemitz et Serge Verry.

3. W. Ruttmann, « Peindre avec le temps », trad. A. Pernet, in Y. Beauvais, *Musique Film*, Paris, éd. Scratch / Cinémathèque française, 1986, p. 18.

4. Cette recherche a bénéficié d'une bourse du Fonds de soutien à l'initiative et à la recherche d'Arcadi Ile-de-France.

5. A. Hofmann, *Graphic Design Manual: Principles and Practice*, New York, Van Nostrand Reinhold, 1965.

« Je pars de situations élémentaires :  
“Un point animé et un son”, “Deux  
lignes droites animées et deux sons”.  
J'étoffe ensuite ces explorations, jusqu'à  
travailler sur des flux d'éléments  
graphiques et de grains sonores régis  
par des algorithmes complexes.



© Birech Montigné

partiels...) et ceux de la création graphique animée (variations de forme, de position, de rotation, d'échelle...). Ces études sont organisées par thématiques (le bruit, l'écho, la vibration...). Je pars de situations élémentaires : « Un point animé et un son », « Deux lignes droites animées et deux sons ». J'étoffe ensuite ces explorations, jusqu'à travailler sur des flux d'éléments graphiques et de grains sonores régis par des algorithmes complexes.

Le fond de ces études progressives est toujours noir, et les figures blanches, ce qui favorise, selon moi, l'attention portée au son. Des textes didactiques et des ressources viennent éclairer des fondamentaux plastiques. Une cartographie des termes donne la possibilité d'une navigation plus transversale.

*C&R : Au-delà du développement d'un outil et de procédés interactifs pour la création audiographique, quels sont les autres volets de vos recherches ?*

**B.M. :** Comme l'ont fait des auteurs tels que Michel Chion<sup>6</sup> pour le cinéma, je cherche à dégager une *typologie de relations audio-visuelles*. Pour inspirer et nourrir mes études, j'ai réuni un corpus d'une trentaine de travaux artistiques, d'ouvrages et de logiciels, dans lesquels la mise en scène d'objets primaires est au cœur de la réflexion. Par exemple : les études animées en noir et blanc d'Oskar Fischinger, les œuvres génératives de Manfred Mohr, les carnets de Véra Molnar, *Le catalogue de sons synthétisés par ordinateur* de Jean-Claude Risset, l'*Audiovisual Environment Suite (AVES)* de Golan Levin, les recherches de Pierre Schaeffer, le logiciel Iannix, les structures de déphasage chez Steve Reich, le logiciel de composition audiovisuelle Whitney-Reed RDTD imaginé par John Whitney.

L'aspect cognitif tient aussi une part importante dans ma recherche. Les lois de groupement entre des formes par proximité, similarité ou continuité, qui constituent la base de la théorie de la Gestalt<sup>7</sup> dans le domaine visuel, peuvent trouver des équivalents dans le domaine sonore. L'analyse de scènes auditives, développée notamment par Albert Bregman<sup>8</sup> à la fin des années 1980, en constitue une bonne illustration, avec des recherches sur la manière dont des sons sont fusionnés ou séparés par le cerveau.

Compte tenu de ces similitudes, de plus en plus de chercheurs adoptent une approche transversale et s'intéressent à la perception multimodale. L'analyse de ces influences perceptuelles réciproques sont autant de pistes me permettant d'envisager des recherches fondamentales aux frontières de l'art et de la science. ■

Étude de densités de bruits graphiques et sonores.

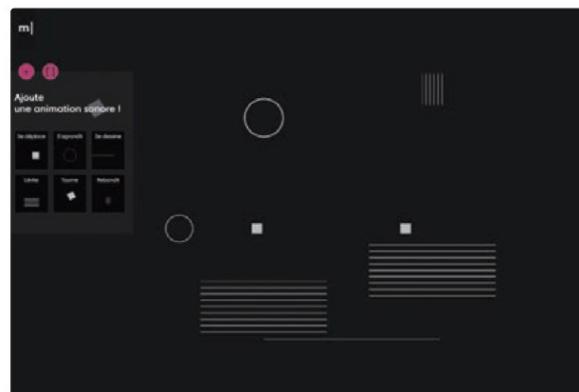
6. Notamment dans son ouvrage *L'audiovision*, Paris, Nathan Université, 1991.

7. *Psychologie de la forme*, théorie née en Allemagne au début du XX<sup>e</sup> siècle.

8. A. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*.

Une version jeune public de l'application Web a été présentée au Centre Pompidou en mars 2018 (voir fig ci-dessous).

[www.audiographiclab.com/musigraphicode-presentation](http://www.audiographiclab.com/musigraphicode-presentation)



# Chercher l'équilibre : dialogues contemporains entre image et son

par *BENOÎT MONTIGNÉ*

Sur la scène, deux artistes manipulent dans l'obscurité leurs machines respectives. L'un est plasticien, l'autre musicien. Le premier projette sur l'écran noir des essais constitués de minuscules points blancs animés qui, tour à tour, se réunissent, se défont et se réagencent pour venir former des nuages géométriques instables et éphémères. Le second génère en temps réel des vagues de bourdonnements et de déchirures électro-acoustiques, en tension avec ces particules en perpétuelle agitation. C'est ainsi que dans *Tempest* d'Antoine Schmitt et de Franck Vigroux, flux visuels et flux sonores s'entremêlent dans des chorégraphies mi-préparées, mi-improvisées, sollicitant notre double perception de spectateur-auditeur.

Ce type de production scénique intégrant de l'image et du son n'est pas nouveau, loin de là. Appelé communément « performance audiovisuelle », il repose sur l'arrivée dans les années 1980 de logiciels informatiques permettant de mixer et de projeter des images afin d'accompagner les performances « live » de musique électronique. Il s'est développé avec l'avènement des festivals numériques, qui ont permis le décroisement de ces pratiques tout en autorisant des démarches plus artistiques. Il n'en reste pas moins que l'image joue souvent un rôle ingrat : illustrer une création à dominante musicale. Inversement, pour une majorité d'œuvres s'inscrivant dans le champ des arts de l'image en

mouvement, une place restreinte est accordée à la création sonore.

Ce n'est toutefois pas le cas dans les expériences proposées dans la programmation 2015-2016 de l'Ensemble intercontemporain. *Tempest*, d'Antoine Schmitt et de Franck Vigroux, *Extension Sauvage* de Jacques Perconte et de Jeff Mills, *Cluster.X* de Kurt Hentschläger et d'Edmund Campion (cf. p. 34), et enfin *No More Masterpieces* du collectif 33<sup>1/3</sup> (cf. p. 66) sont les espaces d'une recherche d'unité, voire de « synesthésie » entre l'image et le son. Au-delà du propos artistique, ces artistes cherchent également à transcender la question de la distinction traditionnelle entre musique et image en mouvement, en donnant naissance à une véritable entité composite, dont chaque « aile »<sup>1</sup> – le visuel, le sonore – résonne avec une force égale.

Dans l'histoire des arts de l'écran, un moment déterminant dans cette quête d'équilibre peut être identifié au début des années 1940 : John Whitney et son frère James réalisent cinq courts métrages d'animation appelés *Five Abstract Film exercises* (1943-1944). Ceux-ci sont conçus en réaction à d'autres travaux antérieurs, qui, s'ils sont tout à fait remarquables (citons en particulier les films du pionnier Oskar Fischinger), partent d'une musique préexistante sur laquelle ils viennent synchroniser des formes animées. Dans la perspective de dépasser ce procédé, les frères Whitney développent non

<sup>1</sup> Pour reprendre l'expression du lettriste Isidore Isou dans son film *Traité de bave et d'éternité* (1951), dont le texte a été édité aux éditions Hors Commerce en 2000.

seulement de manière expérimentale et originale la partie visuelle – à l'aide de pochoirs et de gélatine colorée –, mais mettent également au point un système complexe de pendules leur permettant de synchroniser le mouvement des éléments visuels abstraits avec du son optique. Cette technique alors récente permettait de générer des sons « synthétiques », et cela bien avant la musique électronique et la musique concrète. Par exemple, tandis qu'une forme abstraite s'éloigne dans l'arrière-plan, un son « venu de nulle part »<sup>2</sup> qui lui est associé passe progressivement de l'aigu au grave puis s'estompe. Les Whitney explorent, à travers ces petits films, un certain nombre de relations entre image et son, en veillant en permanence à ce qu'aucune des deux composantes ne prennent le pas sur l'autre. C'est, à notre sens, la première manifestation d'une œuvre réellement audio-visuelle. Les deux frères écrivaient déjà à ce propos en 1944 : « [...] de réelles possibilités créatrices émergent quand la structure de l'image dicte ou "inspire" la structure sonore et vice-versa, ou quand leur conception est simultanée. Ceci évidemment est mieux réalisé quand les deux parties ont des origines créatrices communes »<sup>3</sup>. Contrairement à son frère James qui va au fil des années abandonner ses expérimentations, John Whitney consacrera sa vie à la création. Ce dernier réalisera dans les années 1960 le premier film d'animation artistique par ordinateur, puis développera ses théories sur la complémentarité audio-visuelle<sup>4</sup>. Il mettra celles-ci en pratique grâce à un instrument logiciel, le « Whitney-Reed RDTD », outil de composition audiovisuelle en temps réel qu'il aura lui-même imaginé. Il est donc l'une des figures pionnières de ce que l'on pourrait appeler l'« art audiovisuel »<sup>5</sup>.

Les abstractions « audiographiques » d'Antoine Schmitt et de Franck Vigroux renvoient assez directement aux expérimentations de John Whitney.

Dans *Extension sauvage*, de manière différente mais tout aussi intéressante, ce sont les paysages vidéo de Jacques Perconte, tournés à Madère, qui vont dialoguer avec les compositions minimalistes de Jeff Mills pour tenter d'accomplir une fusion audiovisuelle. Par un traitement complexe de l'algorithme de compression numérique, Jacques Perconte applique à l'image vidéo des pixellisations graduelles qui l'emportent dans des abstractions picturales aux couleurs surprenantes et saturées pendant que Jeff Mills, explorateur familier des rapports entre musique et image, imprime au flux de la vidéo de subtiles oscillations rythmiques et musicales. Dominique Willoughby écrivait à propos du cinéma de Len Lye – un cinéaste d'animation contemporain de John Whitney – qu'il avait « [...] sans doute créé un des meilleurs modèles du cinéma graphique musical, par le choix de ses musiques, la puissance de ses images, et par une dialectique de leurs relations, faites de jonctions et de disjonctions, de synchronisme et d'asynchronisme, ménageant les espaces nécessaires à leurs existences respectives. Le cinéma graphique et la musique ne s'illustrent pas l'un l'autre ni ne s'accompagnent mutuellement, ils se rencontrent et rentrent en résonance »<sup>6</sup>. Cette description s'applique très bien à la rencontre entre les variations électroniques de Jeff Mills et l'art vidéo de Jacques Perconte.

Une autre collaboration, *Cluster.X*, verra s'entre-croiser une vidéo générative de Kurt Hentschläger, l'un des membres du duo Granular-Synthesis (acteur référent de la performance audiovisuelle artistique au début des années 1990), et les compositions contemporaines d'Edmund Campion. On pourra y voir et y entendre d'évanescents silhouettes abstraites, baignées dans une atmosphère sonore dense et énigmatique, se matérialiser progressivement dans un lent mouvement. Si le matériau visuel et sonore est en partie préparé à l'avance, il est diffusé

<sup>2</sup> Thomas Y. Levin, « "Des sons venus de nulle part." Rudolf Pfenninger et l'archéologie du son synthétique », dans le catalogue *Sons & Lumières. Une histoire du son dans l'art du xx<sup>e</sup> siècle*, Paris, Centre Pompidou, 2004, p. 51.

<sup>3</sup> John Whitney, « Musique audio-visuelle : musique couleur – cinéma abstrait », trad. Alain Alcide Sudre, dans Yann Beauvais, *Musique Film*, Paris, Scratch / Cinémathèque Française, 1986, p. 26.

<sup>4</sup> John Whitney consignera ses recherches dans son ouvrage *Digital Harmony : On the Complementarity of Music and Visual Art*, New Hampshire, McGraw-Hill, Peterborough, 1980.

<sup>5</sup> Le terme « audiovisual art » est plus courant en anglais et couvre l'ensemble des œuvres définies par leur rapport entre image et son.

<sup>6</sup> Dominique Willoughby, *Le Cinéma graphique : une histoire des dessins animés, des jouets d'optique au cinéma numérique*, Paris, Textuel, 2009, p. 170.

## DOSSIER Les nouvelles scènes de la musique

et interprété avec un certain degré d'indétermination, le compositeur pouvant notamment intervenir en temps réel sur la spatialisation en multicanal des masses sonores. Il devient d'ailleurs de plus en plus fréquent qu'autour des traitements en temps réel de l'image et du son les artistes intègrent des dispositifs visant à explorer les potentialités de la diffusion sonore, procédé dont le compositeur Karlheinz Stockhausen était l'un des pionniers dans les années 1950. Plus rares en revanche sont les manipulations directes des moyens de diffusion visuelle comme l'écran. C'est un défi que tentera de réaliser le collectif 33<sup>1/3</sup>, à travers non plus une spatialisation du son, mais une « spatialisation » de l'image. Dans *No More Masterpieces* (d'après le *Concerto "Séraphin"*), la base musicale est l'interprétation d'une œuvre contemporaine du compositeur Wolfgang Rihm, inspirée par le travail d'Antonin Artaud. Les artistes ont construit à partir de documents liés à ce dernier une création vidéo qui sera projetée sur un écran « cinématique », dont les dimensions et la forme seront variables. Ce dispositif novateur s'inscrit dans les mêmes préoccupations que celles formulées par le cinéaste Sergueï Eisenstein, qui, dans un discours prononcé en 1930, parlait déjà de la reconstruction nécessaire de la salle de cinéma sonore, et imaginait un dispositif d'écran constitué de surfaces verticales et horizontales : « [...] cela nous apporte un nouvel agent gigantesque d'impression : l'assemblage rythmique des différentes formes d'écran, l'attaque de notre champ de perception par des impulsions émotionnelles associées aux variations géométriques et dimensionnelles qui suivent de la succession des dimensions, des proportions et des plans possibles »<sup>7</sup>.

Les œuvres ici évoquées se distinguent autant du « ciné-concert » que des projections illustratives et redondantes d'images qui viennent accompagner l'interprétation d'une composition musicale pré-existante. Dans ces deux cas, il y a dès le départ

un déséquilibre causé par la force de la composition originale, en particulier si l'œuvre est déjà bien connue par le public (avec l'exception de l'opéra, qui nous a accoutumés à voir différentes mises en scène d'une même composition). Les démarches et collaborations plus innovantes et s'inscrivant dans une recherche véritablement audio-visuelle demandent une attention particulière. À travers un nombre croissant de manifestations et d'expositions, on observe une véritable prise de conscience de l'importance de ces rapports entre les musiques expérimentales, électroniques, électro-acoustiques, contemporaines, les arts sonores, l'art vidéo, l'animation, ou bien encore le cinéma expérimental. Les écoles d'art et les universités commencent à offrir une place réelle à l'enseignement pratique et théorique des arts sonores. Par extension, la pédagogie et la recherche dans les domaines croisés des arts visuels et sonores doit également pouvoir se développer, au-delà des ruptures technologiques et esthétiques entre l'image optique, vidéo et numérique, ainsi qu'entre le son acoustique ou électronique. Des échanges constants et denses entre ces champs artistiques existent dans l'histoire de l'art. Ils viennent éclairer les pratiques artistiques contemporaines utilisant l'image et le son comme matériau et préfigurent les futurs environnements et dispositifs de l'art audiovisuel.

Benoît Montigné est enseignant en multimédia à l'École nationale Supérieure des Arts Décoratifs de Paris (ENSAD), concepteur multimédia pour des institutions culturelles, doctorant en esthétique, sciences et technologie des arts à l'Université de Paris 8. Il est aussi fondateur du site Sonore Visuel, consacré à l'histoire et à l'actualité des arts sonores et audiovisuels.

<sup>7</sup> Sergueï M. Eisenstein, *Le Carré dynamique*, Paris, Séguier, coll. Carré ciné, 1995, p. 44.