

## **Titre de l'article :**

Le clip comme forme d'expression musico-visuelle : pour une esthétique de la relation musique-images

## **Auteur :**

Antoine Gaudin  
Maître de conférences en Études cinématographiques et audiovisuelles  
Université Paris 3 / IRCAV

## **Résumé :**

Cet article propose une réflexion sur la spécificité expressive du vidéoclip musical. À partir de l'analyse de quatre clips, il déploie les enjeux liés à la « primauté » chronologique et compositionnelle de la musique sur les images à l'intérieur de ce secteur de création. Au-delà de sa contribution sur le plan épistémologique (comment pourrait-on approfondir l'approche esthétique de l'objet-clip ?), il vise ainsi une meilleure compréhension de l'expérience sensible et signifiante que les clips font vivre à leurs audio-spectateurs. Il encourage donc à ne pas analyser les clips seulement à l'aide des catégories mobilisées pour les autres formes audiovisuelles (cinéma, télévision, etc.), mais aussi et surtout à adopter une méthodologie d'ensemble respectueuse de leur singularité musico-visuelle.

This article is based on two complementary definitions of the music video : as a media and cultural format on the one hand ; and as a principle of musical and visual creation on the other hand. The main hypothesis of our paper is that we can't fully understand the global phenomenon of the music video if we fail to raise the issue of its expressive singularity – that is, if we continue to consider music video as an audiovisual message as any another, to reduce it to its semantic content and to ignore the particular aesthetic experience it provides to its audio-viewers: that of a specific association between music and images.

## Corps de l'article :

En tant que composition d'images superposée à un morceau musical préexistant afin d'en assurer la promotion auprès d'un canal de diffusion mass-médiatique, le vidéoclip constitue un objet central dans le vaste univers audiovisuel contemporain<sup>1</sup> – dont il reflète et synthétise la plupart des grands enjeux sociaux et figuratifs. Et dans le même temps, il constitue un secteur de création « expérimental-pop » qui prolonge, au cœur même du système industriel de production des objets culturels, une histoire des relations entre le sonore musical et le visible : une histoire au sein de laquelle la plastique musicale, *venant en premier*, impose aux formes visuelles ses principes de composition<sup>2</sup>.

À partir de l'analyse de plusieurs clips bâtis sur une « idée esthétique » pleinement musico-visuelle, il s'agira ici de proposer une réflexion exploratoire sur cette singularité expressive. Ces clips ne seront donc pas considérés comme des messages audiovisuels « comme les autres », et leur étude ne se limitera pas à leur contenu visuel-sémantique ou à leur fonction communicationnelle. En insistant plutôt sur l'« expérience esthétique »<sup>3</sup> particulière qu'ils proposent à leurs audio-spectateurs, on essaiera de montrer que ces réalisations sont susceptibles de transmettre des sensations et des significations spécifiques, irréductibles à tout autre moyen d'expression (du moins, à l'intérieur des grands canaux de diffusion).

D'une définition fonctionnelle et intensive du clip en tant qu'objet médiatique et culturel (celle que formulent les deux premières lignes de cet article), on passe donc à une définition structurelle et extensive du clip en tant que principe de création musico-visuel.

La structure des clips que nous avons retenue ne fait en effet qu'exemplifier, de façon sans doute plus nette et approfondie, ce que l'on retrouve dans la quasi-totalité des clips aujourd'hui : la musique y constitue « la source du mouvement des images ». Par cette expression, on désigne une relation de *primauté* de la musique sur les images, très peu courante dans le monde audiovisuel – qui en général (hors le cas d'une captation expresse d'un numéro musical), privilégie l'image, et relègue la musique à un rôle d'accompagnement<sup>4</sup>.

Cette primauté opère à trois niveaux : au niveau *génétique*, le morceau musical préexiste, il constitue en lui-même une œuvre autonome et il est, la plupart du temps, respecté dans son intégrité (et joué dans son intégralité) au sein du produit final ; au niveau *compositionnel*, la musique a un caractère matriciel, son rythme, sa tonalité, ses dénnotations, ses qualités de timbre, sa texture, sa structure déterminent de manière étroite la création des images (si elle peut prendre des formes diverses, l'« harmonie » de la musique et des images est une règle quasi-absolue dans le clip) ; au niveau *culturel*, enfin, et bien que cet aspect crucial n'ait pas forcément d'effet perceptif direct et avéré (nous pouvons ignorer, en regardant une œuvre audio-visuelle, si elle a été pensée en fonction de la musique ou si c'est la musique qui a été ajoutée aux images), notre savoir latent sur le clip, son statut, sa fonction et la place qu'il occupe dans la filière industrielle du disque et dans les pratiques de

---

<sup>1</sup> À la fin de l'année 2015, les dix plus grandes audiences du canal de diffusion Youtube depuis sa création (dépassant le milliard de vues) étaient des clips. Source : LePoint.fr, « Les dix vidéos qui ont fait exploser les compteurs de Youtube », 28/12/2015, [http://www.lepoint.fr/musique/les-10-vidéos-qui-ont-fait-exploser-les-compteurs-de-youtube-28-12-2015-2005521\\_38.php](http://www.lepoint.fr/musique/les-10-vidéos-qui-ont-fait-exploser-les-compteurs-de-youtube-28-12-2015-2005521_38.php)

<sup>2</sup> Sur le lien du clip à cette généalogie, passant notamment par les cinémas expérimentaux et d'avant-garde et l'art vidéo, voir Brenez (2007) et London (2010).

<sup>3</sup> Au sens d'un mode organique d'être-en-relation au sein d'un cadre quotidien (voir Dewey, 1934).

<sup>4</sup> « Most music we hear in films and television is background music. That is, the images and story exist first, then music is added very late in the production, actually during the "post-production" stage in most cases. » (Burns & Thompson, 1987)

consommation culturelle, conduisent à recevoir les œuvres sous l'angle de la subordination de la bande-image à la bande-musique, et les images comme « conséquence » des morceaux musicaux qu'elles illustrent.

Pour résumer, dans le clip la musique n'est pas convoquée comme accompagnement des images, ce sont au contraire les images qui sont données comme découlant de la musique. Or, bien que cet aspect ait des conséquences importantes sur l'effet (à la fois cognitif et physiologique) produit sur les audio-spectateurs, il a rarement été pris en compte au sein de la littérature théorique et critique sur les clips.

Au sein de cette littérature, les approches esthétiques ne sont pas légion (on mentionnera l'exception notable de Vernallis, 2004), surtout au regard des contributions beaucoup plus nombreuses qui considèrent le clip avant tout comme un objet médiatique et un fait social. En général, les travaux sur le clip demeurent tributaires des catégories classiques d'analyse des *film studies*, focalisées sur l'étude du contenu et de la structure des images et du son sémantique<sup>5</sup>. Ces catégories reflètent : notre hiérarchie sensorielle spontanée, qui fait de la vue et du discours verbal les sources principales d'informations sur le monde ; la domination du visuel et du sonore sémantique telle qu'elle opère dans les études cinématographiques et audiovisuelles ; enfin, la plus grande évidence des catégories analytiques construites pour l'étude des phénomènes visuels ou langagiers, les phénomènes bruitistes ou musicaux paraissant (pour les non-musicologues) « résister » davantage à la démarche de l'analyse verbale rationnelle<sup>6</sup>.

Il est ainsi fréquent que les discours analytiques sur les clips ne s'intéressent pas à la musique, ou qu'ils limitent sur ce point leur propos à des considérations générales sur la pop culture, voire au repérage de correspondances simples entre bande-musique et bande-images (points de synchronisation verticale, lien entre l'« atmosphère » de l'image et la « tonalité » de la musique, etc.) – tout en reposant implicitement sur le présupposé d'une coexistence entre musique et images (tendant éventuellement vers la façon dont les images du clip impactent l'écoute de la chanson), et non d'images qui « découleraient » de la musique, qui seraient modelées par elle<sup>7</sup>. Dans le prolongement des observations effectuées par Carol Vernallis, on pourrait dire à cet égard qu'est souvent appliqué un paradigme aristotélicien (étude du récit, des mécanismes causals, de la construction visuelle et verbale du sens) à une forme musico-visuelle dont les principaux ressorts dynamiques se situent plutôt, selon nous,

---

<sup>5</sup> Il faut aussi dire que dans la période pré-Internet, en particulier les années 80, les travaux existants ont pris appui, par nécessité sans doute autant que par choix, sur les clips les plus diffusés sur les chaînes de TV – qui sont aussi, souvent, les mieux adaptés aux régimes spectaculaires de l'audiovisuel de grande consommation. Cela a pu entraîner, dans les analyses, une focalisation sur des questionnements (la star et sa représentation, l'énonciation du discours, la construction du « récit », etc.) qui constituent des transpositions de ceux que l'on mobilisait alors pour l'étude des autres formes audiovisuelles ; et cela a pu entraîner, par corollaire, une sous-évaluation du clip en tant que mode d'expression autonome. Bien qu'il constitue un étape importante dans la prise en compte du système musique-images dans les clips, l'ouvrage *Experiencing Music Video*, publié avant la création de Youtube, est freiné dans cette voie théorique par le corpus de clips relativement normé avec lequel Carol Vernallis (2004) est encore contrainte de travailler, et qui la conduit à prendre en compte le système musique-images notamment à travers des questionnements tels que l'illustration des paroles de la chanson, l'impact de la structure couplet-refrain ou la mise en valeur de la « star ». C'est donc en dirigeant l'analyse vers des clips issus de cultures musicales plus marginales (musique electro dénuée de paroles et issue d'une culture s'étant construite à rebours du star-system de la pop, notamment) et formellement plus radicaux, que nous tentons, dans cet article, de poursuivre et d'approfondir cette voie prometteuse.

<sup>6</sup> « One difficulty in music video studies (...) is the fact that theorists of the visuals are often not trained in musical analysis and perhaps are not even particularly familiar with or interested in popular music itself. » (Beebe & Middleton, 2007, p.5) Voir également l'article de Christoph Jacke (2010).

<sup>7</sup> Voir Goodwin (1993), qui propose, pour remédier à cet état de fait, l'ébauche d'une « musicologie de la bande-images ». Il prend ici en compte le fait que c'est l'image qui « accompagne » la musique (et non l'inverse), acquérant par là-même certaines des qualités phénoménales de cette dernière.

dans un paradigme héraclitéen : celui d'un flux sculpté par le *rhuthmos*, par sa propre modalité d'écoulement, et analysable au moyen de notions dynamiques-abstraites en provenance directe de l'univers musical, telles que le rythme, le mouvement et la variation.

Certes, le clip musical sollicite autant nos récepteurs visuels que nos récepteurs auditifs, mais nous pensons qu'au niveau de notre système perceptivo-cognitif, il s'adresse au fond moins à notre centre visuel-organisationnel (en quête de cohérence spatiale et temporelle, de solidité et de stabilité du monde représenté) qu'à notre centre auditif-cinétique (en quête de sensations mouvantes et englobantes, moins attaché à un quelconque environnement référent) – et que la sollicitation qu'il effectue ainsi de notre *corps* d'audio-spectateur est une part essentielle de sa poésie singulière.

L'approche que l'on défend ici fait donc des clips les supports d'une expérience esthétique qui ne se limite pas à ce qu'ils *disent* ou *représentent*, mais qui vise avant tout ce qu'ils *transmettent* ou *font sentir*, du fait de leur principe spécifique d'association entre la musique et les images. Pour que cette dimension-là apparaisse pleinement, il faut peut-être porter notre attention sur des œuvres qui proposent autre chose que la performance visuelle de l'artiste musical en situation d'interprétation (clip « performanciel »), ou que l'imitation sur format court du modèle fictionnel cinématographique (clip « narratif »), et qui tendent davantage vers les puissances propres de la composition musico-visuelle.

Dans cette démarche analytique, la difficulté majeure concerne le relatif manque de « prises » théoriques existantes, à l'heure actuelle, sur l'esthétique des musiques actuelles<sup>8</sup>. L'importation de concepts musicologiques pour l'essentiel forgés pour l'étude de la musique classique (instrumentale et écrite) résout mal ce problème, les principes de composition des musiques actuelles (rock, hip-hop, electro, metal, etc.) appelant souvent d'autres catégories d'analyse. Aussi est-il préférable de préciser d'emblée qu'il ne s'agira pas ici de régler toutes les questions qui se posent au niveau de la prise en compte du système musique-images dans les clips, mais plutôt de proposer certaines pistes analytiques d'usage simple (ne requérant pas une expertise musicologique) pour les traiter. À partir de cette ambition théorique raisonnable, on se concentrera essentiellement sur la spécificité expressive du clip, c'est-à-dire sur des expériences sensibles, émotives et intellectuelles qu'en vertu de sa singularité musico-visuelle il est quasiment le seul à prendre en charge au sein des mass-médias contemporains.

Commençons par examiner le clip de *Dirge*, réalisé par Richard Fearless pour le groupe Death in Vegas (1998)<sup>9</sup>. Au niveau strictement visuel, la vidéo s'apparente à une sorte de « mémorial » conceptuel et épuré : elle fait se succéder, sur fond noir, le nom, la photographie et les circonstances de la mort (résumées en quelques lignes) de chacune des 76 victimes ayant succombé à une arme à feu aux Etats-Unis en une seule journée, un certain mardi 2 mai. L'année n'est pas précisée, mais les photos et les quelques lignes de description montées successivement dans le clip viennent d'un dossier intitulé « 7 deadly days », publié par *Time Magazine* en mai 1989.

Étant donné le coût humain élevé de la tolérante législation américaine en matière de port d'armes, on mesure sans peine le poids politique et émotionnel d'un tel matériau-images. Qu'apporte ici, en propre, le principe musico-visuel de composition ? Pour le comprendre, il faut porter notre attention sur la structure du morceau musical. Son principe est simple, mais éclairant pour l'analyse : toutes les huit mesures, le morceau incorpore une ligne instrumentale supplémentaire à son orchestration, les lignes précédemment posées continuant

---

<sup>8</sup> Sur cette question, voir les ouvrages récents de Philip Tagg (2013) ou de Christophe Pirenne (2012).

<sup>9</sup> Toutes les vidéos citées dans cet article peuvent être consultées très facilement sur Youtube. En ce qui concerne *Dirge*, ne pas confondre avec la seconde vidéo officielle, réalisée par Richard Fenwick & Richard Fearless à partir d'une version remixée et plus courte de la chanson. La version que nous analysons est donc réalisée par Richard Fearless seul (Fearless est par ailleurs un des membres fondateurs du groupe Death in Vegas).

à tourner en boucle, sur le même leitmotiv mélodique. Parallèlement, à l'image, un fait divers succède à un autre fait divers, au rythme d'un nouveau tous les huit temps (équivalent à deux mesures).

On peut ainsi décrire un premier effet musico-visuel, assez évident : l'accentuation de la régularité du défilement. Seul, l'œil évalue mal les intervalles de temps ; c'est le concours de l'oreille, qui, en s'installant dans le rythme du morceau musical, informe notre cerveau de l'implacable régularité avec laquelle les photographies des victimes se succèdent les unes aux autres. Or, cette *régularité* de la mort par balles constitue un véritable enjeu politique et esthétique ; associée à la tonalité émotive impassible et répétitive du morceau musical, et à son absence de développement mélodique (qui confère au clip la dimension désabusée du constat), cette régularité devient assez rapidement un phénomène scandaleux (la mort par balles devrait être *exceptionnelle*, et non *régulière* comme ici) et inarrêtable : notre corps d'audio-spectateur comprend très vite qu'à la fin de chaque hypermètre de huit temps une nouvelle victime apparaîtra à l'écran, et il s'installe dans cette litanie périodique. La composition musico-visuelle est donc essentielle pour la perception de la régularité du défilement des images et pour les valeurs affectives et signifiantes que nous lui attribuons.

Un second effet de la composition musico-visuelle se révèle plus décisif encore. Revenons au principe de construction musical : toutes les huit mesures, une nouvelle ligne instrumentale est ajoutée à l'orchestration, les autres lignes précédemment posées continuant à se répéter, en boucle. Cela implique que l'orchestration musicale ne cesse de se densifier (en termes d'occupation du spectre de fréquences, on pourrait dire : de « gonfler ») pendant les cinq minutes du clip. Si, sur le plan visuel, l'apparition d'une nouvelle victime (et de son histoire) entraîne la disparition de la précédente, sur le plan sonore, en revanche, les lignes instrumentales successives s'ajoutent les unes aux autres, pendant que l'orchestration du morceau prend de plus en plus d'ampleur (et que l'espace disponible dans le spectre sonore se raréfie). Autrement dit, sur le plan purement visuel (si l'on jouait le clip sans le son), les victimes ne feraient que se *succéder* ; tandis que, sur le plan de la composition musico-visuelle, l'association de nos perceptions, guidée par l'ajout permanent de nouvelles lignes instrumentales, nous communique l'impression qu'elles *s'empilent* les unes sur les autres. Or, un *empilement* de victimes, cela renvoie à l'idée d'un charnier, qui est beaucoup plus forte et significative qu'une simple *succession* de faits divers isolés, séparés géographiquement aux quatre coins d'un vaste pays. C'est bien dans le corps musico-visuel de l'œuvre qu'est inscrite cette idée sensible de l'empilement, et l'indignation légitime qui l'accompagne chez l'audio-spectateur peut alors « gonfler » en même temps que l'orchestration musicale du morceau. Pour analyser ce phénomène dans le clip, il faut donc avoir *écouté* en priorité, sans pour autant avoir cessé de *regarder*.

Ainsi, en faisant partir l'analyse d'une poignée de catégories musicales (la métrique et l'orchestration) dont le fonctionnement basique est repérable par tous, il est possible d'approfondir substantiellement la compréhension de l'impact sensible et signifiant du clip. Nul besoin, donc, de se transformer en musicologue accompli, pour prendre en compte l'engagement primordial de notre corps d'audio-spectateur dans cette structure musico-visuelle.

L'idée que la sollicitation esthétique d'un clip prend sa source dans la stimulation du corps spectatorial n'est pas neuve en soi. Devant un clip, notre corps se trouve assez naturellement « entraîné » par la combinaison des rythmes du morceau musical et du montage des images sur les points de synchronisation (de coupure comme de convergence<sup>10</sup>). Cet effet physiologique de base a cependant été moins souvent étudié pour ses potentialités expressives

---

<sup>10</sup> Sur ces catégories d'analyse, voir Michel Chion (1990, p.52).

que dénoncé (comme une « formule » simpliste, racoleuse, voire manipulatrice) à l'intérieur de discours dépréciateurs au sujet du clip<sup>11</sup>. Or, lorsqu'on sort des chorégraphies normées des superstars de la pop, on s'aperçoit que cet effet peut aussi bien être mobilisé pour intensifier une idée esthétique qui va par exemple se nourrir de lui afin de le questionner en retour.

Ce mécanisme est à l'œuvre dans le clip du morceau *Rocker*, du groupe electro allemand Alter Ego, réalisé par la vidéaste Corine Stübi (2004). Comme dans d'autres travaux de Stübi (par exemple sa vidéo *Working Girls*), le sujet de la vidéo est le corps féminin, ou plutôt ses stéréotypes contemporains, tels qu'ils investissent les productions audiovisuelles grand public. Visuellement, le clip propose un défilé de « rôles de femmes » archétypaux (le mannequin, la secrétaire, la femme au foyer, la serveuse, la star glamour, etc.) aux mouvements robotiques, portraituretés dans des attitudes de séduction superficielles dont l'érotisme éthéré vient en droite ligne des représentations courantes véhiculées par la publicité, les magazines, les fictions audiovisuelles, et bien sûr... les clips eux-mêmes<sup>12</sup>. La dimension autoréflexive de l'œuvre est patente, notamment dans sa partie finale, qui reproduit de façon décalée un cliché du clip de grande consommation : deux jeunes femmes richement parées pour une sortie nocturne dansent et consomment du champagne dans une limousine (le saut moussant du liquide hors des bouteilles composant, comme dans certains clips de gangsta rap, un symbole phallique massif d'évidence et de vulgarité). Dans la même perspective, le clin d'œil final qui nous est adressé en regard-caméra par l'une de ces femmes, interpelle le spectateur et « teste » son degré d'adhésion à ce qu'il vient de visionner.

On constate, comme dans le cas de *Dirge*, que le matériau-images semble déjà assez opérant par lui-même, dans sa dénonciation de certaines représentations oppressantes de « l'idéal féminin » au sein des médias grands publics. Mais là encore, il est possible d'approfondir l'analyse, en « écoutant en premier ». On remarque alors que la réalisatrice exploite de façon systématique les possibilités de synchronisation et de synchrèse<sup>13</sup> entre la musique et les images. Cette concordance extrême concerne à la fois les coupes du montage (synchronisation des *cuts* visuels sur le *beat* musical) et le mouvement des corps (chorégraphie mécanique des modèles féminins), mais également la matière-image elle-même (quand un *scratch* se fait entendre dans le morceau, on voit à l'image une brève réversion avant retour à vitesse normale de défilement). Stübi peut ainsi modeler les corps de ses modèles féminins dans un rapport d'asservissement total aux dynamiques du morceau musical : les *jump-cuts* sur des fragments de leur physionomie (jambes, visage, mains, poitrine) se multiplient, synchronisés sur les syncopes percussives de la musique ; les mouvements des corps-modèles semblent produire eux-mêmes les bruits aigus de pistons qui accompagnent une bonne partie du morceau (l'audio-spectateur a alors l'impression que ces corps-machines sont animés par des rouages) ; les effets de réversion à l'image (responsables de l'aspect robotique des mouvements) accompagnent systématiquement les *glissandi* musicaux, dont la tonalité dissonante et synthétique produit un effet de *rewind* sonore ; etc. Ainsi « embaumés » rythmiquement dans l'accomplissement méthodique de gestes quotidiens (mettre du rouge à lèvres, enfiler un bas, servir le café, etc.), ces corps de femmes livrées aux regards ressemblent à de parfaits automates. On comprend alors qu'au lieu de se contenter de

---

<sup>11</sup> Voir les exemples de citations amassées par Péquignot (2010).

<sup>12</sup> Sur la fragmentation et la déshumanisation des corps dans la « grammaire » esthétique du clip, voir Burns & Thompson (1987)

<sup>13</sup> On reprend cette notion à Michel Chion, chez lequel il désigne « un phénomène psycho-physiologique spontané et réflexe, universel, dépendant de nos connexions nerveuses, et ne répondant à aucun conditionnement culturel, et qui consiste à percevoir comme un seul et même phénomène se manifestant à la fois visuellement et acoustiquement la concomitance d'un événement sonore ponctuel et d'un événement visuel ponctuel, dès l'instant où ceux-ci se produisent simultanément, et à cette seule condition nécessaire et suffisante » (Chion, 2003, p.192).

« dénoncer » les stratégies spectaculaires déployées autour du corps féminin dans les grands médias, Corinne Stübi a sensiblement amplifié ces dernières, jusqu'au point où nous pouvons éprouver plus directement leur violence et leur absurdité.

Qu'apporte ici la forme musico-visuelle du clip ? Son effet le plus évident est l'amplification, par la mobilisation de tous les moyens visuels et sonores de la composition, de l'idée sensible de mécanisation du corps. À première vue, il semble donc n'être question que d'une intensification de la représentation visuelle ; mais une autre implication plus profonde ne tarde pas à surgir. Car, inscrit dans la matière musico-visuelle de l'œuvre, cet effet de mécanisation atteint directement notre corps d'audio-spectateur. La sollicitation de nos intérocepteurs impliqués dans la sensation projective de la danse, qui opère comme dans la majorité des clips qui synchronisent des mouvements sur de la musique (la plupart du temps au nom du tout-puissant agrément, dans ce qui se voudrait une innocence festive et dénuée d'implications), introduit ici un questionnement politique et sensible de plus grande ampleur : en déployant à l'excès cette stratégie de synchronisation au profit d'une mise en crise du spectacle courant du corps féminin, la réalisatrice questionne la façon dont les mécanismes musico-visuels des industries culturelles, ceux du clip de grande consommation, mais aussi ceux de la télévision ou de la publicité, sollicitent en permanence l'adhésion de notre corps de spectateur à des représentations stéréotypées et dégradantes. À ce niveau d'implication dans le fonctionnement des dites industries culturelles (*Rocker* est un des tubes de l'été 2004, et son clip passait alors en boucle sur MTV), l'opération est plutôt intéressante.

Il n'est par ailleurs pas indifférent que les exemples de *Dirge* et de *Rocker* concernent des clips réalisés à partir de morceaux de musique électronique. Le développement de ce vaste genre musical, devenu vraiment grand public à partir des années 1990, a en effet contribué à libérer les concepteurs de clips des modes de composition « performanciers » (visualisation de la l'artiste musical ou du groupe de musique en situation d'« interprétation ») et « narratifs » (imitation sur format court du modèle fictionnel cinématographique) – qui étaient favorisés par le format « chanson », dans la mesure où ce dernier comportait des paroles signifiantes à illustrer, un interprète-star à magnifier, etc. À l'inverse de ces modes de composition liés aux stratégies spectaculaires du court-métrage de fiction et des shows télévisés, l'electro a globalement permis aux réalisateurs de clips d'affermir la spécificité de leur moyen d'expression en explorant avec davantage d'insistance les possibilités de la composition musico-visuelle<sup>14</sup>.

Les caractéristiques des musiques électroniques (culture anti-star de l'anonymat, absence fréquente de paroles ou de « sujet » de chanson, court-circuitage de l'idée traditionnelle de « performance » instrumentale, expériences singulières de rythme, de tonalité et de texture sonore, etc.) ouvrent ainsi un champ de possibles, émancipé notamment de la contrainte du « sujet » à traiter, ou de la représentation promotionnelle de l'artiste musical en situation d'interprétation. Par ailleurs, au niveau esthétique, les musiques électroniques proposent des structures spécifiques au niveau temporel (l'implacable régularité machinique du *beat*, sans commune mesure avec celle d'un batteur humain) et spatial (les effets d'étendue et de compression des « nappes sonores »), tout en déployant des pratiques (*deejaying*, *sampling*, *mash-up*) et des principes fondamentaux de composition (le fragment, la boucle, la répétition, la superposition, etc.) propres à stimuler l'inventivité des réalisateurs.

---

<sup>14</sup> Un des meilleurs exemples de ce processus créatif est le clip de Michel Gondry pour le morceau *Star Guitar* des Chemical Brothers (2002), dans lequel la recomposition numérique du paysage défilant à la fenêtre du train devient comme la partition rythmique et plastique de la musique – cela, par la synchronisation parfaite entre le passage à l'écran des structures visuelles et l'intervention dans la bande-son des divers fragments musicaux qui composent le morceau. Voir Gabrielli (2004).

Si c'est bien essentiellement au contact de l'électro que le clip a pu, tout au long des années 1990 et 2000<sup>15</sup>, approfondir son esthétique de pure partition rythmique et plastique, il faut également noter que cette tendance a ensuite pu toucher des vidéos tournées pour des genres musicaux plus traditionnels (le rock, la pop, etc.).

Prenons l'exemple du clip réalisé par Martin de Thurah pour le groupe Carpark North, illustrant un morceau, *Human* (2006)<sup>16</sup>, reposant sur un schéma harmonique et structurel (couplet-refrain) relativement courant. Le sujet de cette vidéo (sans lien direct avec les paroles de la chanson, qui évoquent de façon classique une rupture amoureuse) est le corps adolescent, plus exactement le passage de l'enfance à l'adolescence, lorsque frappent les premiers signes de la puberté. On conviendra qu'il s'agit d'un thème assez présent dans les domaines de l'audiovisuel grand public (cinéma de fiction, télévision, etc.). Il constitue donc un objet de choix pour montrer en quoi le traitement musico-visuel du clip peut donner accès à des dimensions sensibles qui sont largement inaccessibles aux formes plus traditionnelles de représentation.

On peut commencer par remarquer que le morceau de Carpark North possède une dynamique rythmique assez nerveuse et syncopée, et que c'est cet aspect que le réalisateur a retenu pour modeler sa bande-image. Il utilise notamment la caméra tremblée, le montage accéléré (incorporant notamment, sous formes d'inserts, des vues microscopiques de cellules, d'artères et de synapses symbolisant l'activité hormonale) et les variations de vitesse de l'image (et tous les trucages « naturels » qu'elles permettent), afin de figurer, par cet alliage musico-visuel très organique en lui-même, la secousse énergétique, le trauma somatique de la puberté affectant les corps des enfants successivement filmés.

Les premières images montrent ainsi ces enfants dans une école, chacun occupant seul son espace (lavabos, vestiaire, gymnase, couloir, salle de classe). Dès les premières images, plusieurs phénomènes – le pas de danse syncopé esquissé par une jeune fille, la brusque volte-face d'une autre fille qui à la faveur d'un accéléré semble disparaître sous sa propre chevelure, ou encore l'étrange bouillonnement de l'eau dans la bouteille tenue par un garçon – perturbent la douceur apparente de ces « portraits » de préados. Ces images sont étroitement synchronisées avec l'ouverture du morceau musical, dont l'orchestration réduite (batterie, voix et clavier) et le faible volume préparent les éclats à venir. De fait, les entrées successives des instruments qui complètent l'orchestration et augmentent l'intensité du morceau (guitare basse au bout de quatre mesures, puis guitares électriques, l'une au vingt-septième temps, l'autre au quarante-troisième temps), se répercutent aussitôt à l'image : sculptés dynamiquement par les secousses du cadre, les jump-cuts du montage et les variations de vitesse de l'image, les corps des enfants, soudain engagés dans des chorégraphies instables, se mettent à trembler, à sauter, à se tendre et se cambrer, à haleter, à tournoyer, à bondir, à planer...

Les éléments du décor qui les entourent se font l'écho de ces décharges énergétiques initiées par la musique : par exemple, sur l'entrée de la première guitare (qui provoque un net « pain » sonore), un garçon, lassé des devoirs qu'il effectuait sagement dans les images précédentes, frappe son pupitre du poing, et c'est l'ensemble des meubles de la classe qui tremblent sous l'impact de cette force nouvelle ; au moment du traditionnel « pont » de la chanson, c'est cette fois l'épuration soudaine de l'orchestration (il ne subsiste que le riff du

---

<sup>15</sup> En plus des éléments que je viens d'apporter, je me permets de renvoyer à Gaudin (2015) pour de plus amples développements sur cette question.

<sup>16</sup> Lauréat du Grand Prix du Festival international des Arts du clip 2007, où je l'ai découvert (le groupe Carpark North n'étant pas distribué en France). J'ai eu l'occasion, il y a quelques années, de proposer une analyse de ce clip pour la revue *Vertigo*. Je la reprends ici de façon différente, en saisissant l'occasion d'approfondir la question de son fonctionnement musico-visuel – cette dernière étant restée secondaire dans l'article mentionné, qui se concentrait surtout sur la représentation visuelle des corps adolescents (Gaudin, 2013).



synthé et la voix du chanteur, soumise à un effet de vocoder) qui motive l'estompement de la lumière ambiante, précipitant dans l'obscurité l'espace d'un gymnase où un autre garçon, torse-nu, tend et détend ses muscles face à la caméra, en « respirant » au rythme de la musique ; au moment précis de la transition entre le pont et la seconde occurrence du refrain, un effet sonore d'aspiration est produit au synthé, et la rencontre d'un garçon et d'une fille, lancés à pleine vitesse l'un vers l'autre puis stoppés nets, comme frappés au cœur de leur élan par une puissance invisible, s'effectue sous le régime d'un choc nouveau. Les « portraits » initiaux se transforment ainsi en éclats somatiques, sporadiques et inquiets, qui épousent scrupuleusement les dynamiques du morceau musical.

Par ailleurs, sur les deux occurrences du refrain de la chanson, le déploiement orchestral et mélodique (souligné à l'écran, la première fois, par la course et l'envol soudain d'une jeune fille) amène une série de plans sur fond noir, qui extraient les corps de leur univers scolaire pour les transporter dans un espace indéterminé où la caméra peut, à loisir, traquer les détails révélateurs d'une spectaculaire tension organique. À la merci d'un souffle provoqué par la musique qui, alternativement, soutient leurs gestuelles ou déforme violemment leurs physionomies, les enfants demeurent là encore solitaires à l'image – même si ce « grand huit » somatique les concerne tous.

Cette dynamique musico-visuelle explosive, qui a pris le contrôle de corps ballottés par leur propre croissance, communique de façon intime à notre corps d'audio-spectateur une idée sensible du passage de l'enfance à l'adolescence : le trauma somatique que la puberté déclenche, l'auto-érotisme à la fois euphorique et inquiet qui l'accompagne, les horizons qui s'éveillent à la conscience (les essais de *look* ou de démarches par les préados), tandis que d'autres se ferment (les « rôles » sexués se durcissent, vecteurs d'angoisses et d'interrogations), c'est tout cela, le champ d'expériences inscrit dans la substance musico-visuelle du clip de *Human* – cette même substance que n'investissent pas les formes dominantes de cinéma et de télévision. En ce sens, on peut dire que le mouvement sensible imprimé par la forme musico-visuelle de ce clip sert de socle à une esthétique pragmatique qui recèle une certaine dimension critique, en ce qu'elle permet de penser et de sentir des phénomènes existentiels laissés à la marge par les grands médias traditionnels.

Ce constat est également valable pour la vidéo de *Get Free*, réalisé par SoMe (de son vrai nom : Bertrand Lagros de Langeron, par ailleurs graphiste et directeur artistique du label musical Ed Banger) pour le groupe Major Lazer (2012), qui constituera le dernier exemple analysé dans cet article. Le réalisateur prend ici appui sur le rythme lent et chaloupé, les basses creusés et le spectre sonore épuré du morceau musical d'origine (un dancehall planant) pour composer une véritable « symphonie de mouvements », parfois modelés par des variations de vitesse (ralenti, accéléré). Filmés comme à la volée dans divers quartiers de Kingston, incarnant des instantanés de vie quotidienne ou des fragments de fêtes nocturnes, ces mouvements sont de toute sorte : danseurs, musiciens, boxeurs, joueurs de dominos, motos, animaux, marionnettes, etc., se succèdent à l'écran, sans qu'aucun des corps ou des objets filmés n'occupe plus d'une poignée de plans rapides dans le montage final. Pourtant, une impression de grande continuité est produite par les nombreux raccords qui relient entre eux les membres de ce qui apparaît comme une véritable communauté de gestes et d'attitudes.

Ces raccords de motifs et de mouvements ne peuvent tous être cités ; pêle-mêle, on mentionnera ceux qui relient : l'abattement d'un jeu de dominos sur la table d'un café et la manipulation d'une table de mixage par un DJ ; les « convulsions » d'une marionnette dans la rue et celles d'une jeune femme en pleine *booty dance* dans une boîte de nuit ; le nouage des tresses d'une autre jeune femme dans un salon de coiffure et celui des chaussures d'un boxeur avant de monter sur le ring ; l'arrêt brusque d'une moto et le geste de suspension d'un danseur ; etc. C'est un peu comme si, tout au long du clip, chaque corps, chaque objet

mouvant, chaque phénomène filmé (chacun dans un espace-temps qui lui est irréductiblement propre), se passait le relais d'un vaste mouvement organique commandé par la dynamique et la texture du morceau musical.

Car ce que produit ici l'alliage musico-visuel, au-delà de convaincantes synchronèses ponctuelles (comme lorsque l'eau de mer se retire sur l'entrée dans le morceau d'une harpe électronique), c'est une sensation marquante de suspension et d'aspiration. Traditionnellement inscrite dans la production des morceaux du genre dancehall (par une épuration de l'orchestration mélodique, une amplification des basses, un marquage net des contre-temps et des effets de réverbération), cette sensation trouve un écho visuel dans l'intégration des corps au sein d'une singulière chorégraphie de gestes – saisis pour leur aspect « aérien », alors même que le cadre et la « scénographie » insistent de leur côté sur la lourdeur et la massivité des corps qui les produisent. À mille lieux d'un rapport hédoniste et calibré au corps festif auquel sacrifient souvent les clips de grande consommation, ce qui émerge ici, c'est donc une chorégraphie d'ensemble dont la qualité de mouvement se distingue immédiatement par sa forte identité.

Cette (audio-)vision neuve et incarnée de la Jamaïque et de la vie de ses habitants (aucun cliché touristique ou occidentalisme ici) charrie avec elle, en accord avec la tonalité mélancolique du morceau musical, une ambiguïté et une richesse sensorielles exceptionnelles. Le potentiel synesthésique de cette vaste pulsation musico-visuelle est en effet remarquable : porté par le mouvement sensible inscrit dans son système musique-images, le clip réunit ces fragments de la nuit tropicale (chaleur, fumée, humidité, etc.), tout en extrayant certaines dynamiques essentielles (intensité, indolence, etc.) des blocs de vie collectifs qui la constituent. On le voit (et on l'entend en même temps) : la force du clip, ici, c'est justement de ne pas avoir à raconter d'histoire structurée, ni à « documenter » (au sens didactique du reportage explicatif) ce qu'il montre ; sa force réside, au contraire, dans sa capacité à composer une suite musicale et visuelle, une succession d'espaces, de mouvements de rythmes et d'atmosphères, qui au fond nous en dit peut-être beaucoup plus que n'importe quels « fiction » ou reportage, sur les données sensibles de l'existence dans cet endroit du monde (il nous en dit en tout cas *autre chose*).

Ces quelques exemples nous conduisent au cœur de la spécificité musico-visuelle du vidéoclip en tant que médium autonome. Insistance sur la régularité accumulative des décès par balles aux Etats-Unis (*Dirge*), sollicitation autoréflexive du corps spectatorial afin de tester son adhésion à l'iconographie mass-médiatique du corps féminin (*Rocker*), épreuve du choc somatique de l'adolescence (*Human*), impact sensoriel et énergie collective d'une ville jamaïcaine (*Get Free*) : autant d'expériences esthétiques primordiales qui imposent, pour être perçues et analysées, de prendre en compte, à sa juste mesure, le fonctionnement propre du système musique-images dans le secteur du clip. Sans nécessairement constituer les formes dominantes de ce média, les œuvres analysées nous intéressent parce qu'elles en manifestent les puissances expressives les plus marquantes (au regard des autres formats audiovisuels). Pour en rendre compte, le défi de l'analyse consiste à tirer toutes les conséquences de cette singularité, et à déployer, à partir de la sollicitation du corps de l'audio-spectateur, une *aisthesis* adaptée à cette primauté de la musique sur l'image. Certains clips nous y invitent plus que d'autres, certes, mais il ne faut jamais oublier que le clip participe avant tout au « monde de l'art » des musiques actuelles, et que dans les processus de consommation culturelle, l'exposition aux clips s'effectue prioritairement par le biais des musiques qu'ils illustrent. Ainsi, au-delà de son apport pour l'analyse immanente, l'intérêt de cette proposition d'attention à la forme musico-visuelle des clips pourrait être aussi de mieux comprendre leur fonctionnement dans l'espace public. Ce qui est susceptible d'émerger de cette démarche, c'est une vision du clip qui ne serait plus seulement celle d'un objet médiatique défini par le

contenu de ses images ou de ses « discours », mais bien celle d'un art pop à la fois ludique et critique, d'une forme-flux façonnée par ses propres modalités d'« écoulement ». C'est ainsi que nous pourrions évaluer la portée des expériences esthétiques que propose aujourd'hui cette forme d'art modeste et non sanctuarisée, accessible à tous sans concepts, et porteuse de significations sensibles qui ne sont prises en charge, aujourd'hui, par aucune autre forme audiovisuelle grand public.

### **Bibliographie :**

- BEEBE Roger & MIDDLETON Jason, eds. (2007), *Medium Cool. Music Videos from Soundies to Cellphones*, Durham, Duke University Press, 360 p.
- BRENEZ Nicole (2008), « Clip, pamphlet, subversion. Jalons pour une histoire de la contre-culture populaire », in *Playback*, catalogue de l'exposition au Musée d'Art Moderne de la Ville de Paris, pp. 19-25.
- BURNS Gary & THOMPSON Robert (1987), « Music, Television and Video: Historical and Aesthetic Considerations », in *Popular Music and Society*, 11 : 3, pp. 11-25.
- CHION Michel (1990), *L'audio-vision*, Paris, Nathan, 240 p.
- CHION Michel (2003), *Un art sonore, le cinéma*, Paris, Cahiers du cinéma, 512 p.
- DEWEY John (1934), *L'art comme expérience*, Paris, Gallimard, 2010, 608 p.
- GABRIELLI Giulia (2010), *An Analysis of the Relation between Music and Image : the Contribution of Michel Gondry*, in Keazor Henry & Wübbena Thorsten (eds.), *Rewind, Play, Fast Forward : The Past, Present and Future of the Music Video*, Bielefeld: Transcript Verlag, 2010, pp. 89-109.
- GAUDIN Antoine (2013), « L'aventure intérieure de l'adolescence : à propos de deux clips de Martin de Thurah », in *Vertigo*, n°45, « L'empire de l'adolescence » ; « Le vidéoclip, laboratoire expérimental de l'industrie culturelle : innovations technologiques et expériences esthétiques à l'ère numérique », in Raphaël Jaudon, Dario Marchiori et Luc Vancheri (dir.), *Ecrans*, « Expanded Cinema », n°3, 2015-1, pp. 163-179.
- GOODWIN Andrew (1993), *Dancing in the Distraction Factory. Music Television and Popular Culture*, New York, Routledge, 237 p.
- JACKE Christoph (2010), « Who Cares about the Music in Music Videos ? Toward a Multiperspectival Pop Culture Study of Music Videos », in Keazor Henry & Wübbena Thorsten (eds.), *Rewind, Play, Fast Forward : The Past, Present and Future of the Music Video*, Bielefeld: Transcript Verlag, pp. 179-194.
- LONDON Barbara, « Looking at Music », in Keazor Henry & Wübbena Thorsten (eds.), *Rewind, Play, Fast Forward : The Past, Present and Future of the Music Video*, Bielefeld: Transcript Verlag, 2010, pp. 59-65.

PEQUIGNOT Julien, « Pour une sociologie méta-interprétative : le clip et ses discours, de la tentation postmoderne à la nécessité pragmatique » (2010), thèse de doctorat de l'Université Lyon 3, sous la direction de Jean-Pierre Esquenazi.

PIRENNE Christophe (2012), *Une histoire musicale du rock*, Paris Fayard, 800 p.

TAGG Philip (2013), *Music's Meaning : a Modern Musicology for non-musos*, New York, Mass Media Music Scholar's Press, 710 p.

VERNALLIS Carol (2004), *Experiencing Music Videos : Aesthetics and Cultural Context*, New York, Columbia University Press, 480 p.

### **Biographie :**

Antoine Gaudin est maître de conférences en Études cinématographiques et audiovisuelles à l'Université Paris 3 – Sorbonne nouvelle. Il est l'auteur de *L'espace cinématographique : esthétique et dramaturgie* (Armand Colin, 2015).

**Mots-clés :** analyse musicale ; clips vidéos ; corps ; expérience ; esthétique ; corps

aesthetics ; analysis (musical) ; body ; experience ; video (music) / clip ;

### **Liste des genres musicaux :**

electro pop /rock ; dancehall / raggamuffin ; indie rock

### **Liste des artistes étudiés :**

Groupes de musique : Death in Vegas ; Alter Ego ; Carpark North ; Major Lazer

Réalisateurs-vidéastes : Richard Fenwick ; Corinne Stübi ; Martin de Thurah ; SoMe