



HAL
open science

Dire la vérité à propos du terrorisme – La socio-logique des vidéoclips de plaidoyer

Baudouin Dupret, Enrique Klaus

► **To cite this version:**

Baudouin Dupret, Enrique Klaus. Dire la vérité à propos du terrorisme – La socio-logique des vidéoclips de plaidoyer. Les Cahiers d'ethnométhodologie, 2010, 4, pp.53-73. <halshs-00577673>

HAL Id: halshs-00577673

<https://shs.hal.science/halshs-00577673v1>

Submitted on 17 Mar 2011

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.



HAL Authorization

DIRE LA VERITE A PROPOS DU TERRORISME

La socio-logique des vidéoclips de plaider

Baudouin DUPRET, baudouindupret@skynet.be

Enrique KLAUS, enrique.klaus@laposte.net

_ Résumé :

Les chaînes de télévision produisent une grande variété de vidéoclips, dont certains ont une tonalité politique de plaider. En regardant des chaînes télévisées arabophones, on peut observer une série de vidéoclips qui, en moins d'une minute, proposent des narrations contrastées dénonçant la terreur et le terrorisme. Dans son attitude naturelle, le téléspectateur comprend ces histoires à l'intérieur d'une grammaire pratique faite de procédés ordinaires de raisonnement, de collections catégorielles partagées et de connaissance d'arrière-plan sur des questions pertinentes pour l'audience dont il est membre. Dans cet article, nous analysons la grammaire pratique que ces clips mettent en place dans le jeu de langage spécifique à la narration de vérité à propos du terrorisme.

Ethnométhodologie — Études des médias — Socio-logique du vidéoclip — *Gestalt* — Terrorisme — Moyen-Orient

_ Abstract :

TV channels produce a whole range of video-clips, some of them with a political, advocacy tonality. Watching Arabic-speaking TV channels, we observed a series of video-clips which, in less than one minute, proposed contrasted narratives denouncing terror and terrorism. The spectator, in his natural attitude, understands these stories within a practical grammar made of ordinary sense-making practices, shared categorization devices, and background knowledge of issues relevant for the audience he is a member of. In this paper, we address the practical grammar these clips displayed of the specific language game of truth-telling narratives about terror.

Ethnomethodology — Media Studies — Socio-logic of video-clip — *Gestalt* — Terrorism — Middle-East

_ Plan d'article :

1. La socio-logique du vidéoclip
2. La grammaire pratique de vidéo-clips contrastifs à propos du terrorisme

Plans

Séquences

Narrations

Conclusion

Les chaînes de télévision produisent une grande variété de vidéoclips, dont certains ont une tonalité politique de plaidoyer. En regardant des chaînes télévisées arabophones, on peut observer une série de vidéoclips qui, en moins d'une minute, proposent des narrations contrastées dénonçant le terrorisme. Dans son attitude naturelle, le téléspectateur comprend ces histoires à l'intérieur d'une grammaire pratique faite de procédés ordinaires de raisonnement, de collections catégorielles partagées et d'une connaissance d'arrière-plan sur des questions pertinentes pour l'audience dont il est membre. Dans cet article, nous analysons la grammaire pratique que ces clips mettent en place dans le jeu de langage spécifique à l'énonciation de vérités sur le terrorisme.

Épistémologiquement, nous nous situons dans le courant de la praxéologie et de la philosophie du langage, pour poursuivre l'exploration de la production d'objets médiatiques et, particulièrement, de la relation qu'elle reflète entre les techniques audiovisuelles et les revendications de véridiction [Dupret et al. 2007 ; 2008]¹. Inspirés par une ethnométhodologie wittgensteinienne, nous posons que ces vidéoclips qui associent des images, des textes, des sons et de la musique ont une intelligibilité naturelle. Ces clips créent en effet une structure de pertinence directement (quoique, peut-être, imparfaitement) disponible à une audience ordinairement compétente de ces médias. Une telle audience comprend ces vidéoclips organisés normativement et moralement par le moyen de pratiques de raisonnement normativement et moralement constituées. Notre but est dès lors de décrire analytiquement la production audiovisuelle disponible pour les membres de l'audience et les méthodes que ces derniers doivent mobiliser pour comprendre la narration et le message de son plaidoyer. Cela montre comment ces vidéoclips produisent et sont produits par un maître récit dont la méthode d'interprétation est documentaire, c'est-à-dire, rétrospectivement implantée dans ce qui est « communément su » et prospectivement orientée vers les inférences ultérieures qu'il rend possibles.

Comme Hutchby le résume [2006, p. 6-10], les études sur les médias se sont longtemps focalisées sur les effets de la radiodiffusion sur son audience. C'est ici que l'on peut situer les débats sur l'« *effet narcotique* » des médias de masse, leur contribution à l'appauvrissement de la culture et à la désensibilisation à la violence. Le principal résultat de ces recherches a été de traiter les membres d'audiences médiatiques comme des « *idiots culturels qui sont conditionnés ou sur lesquels on agit par le média, à l'extérieur de leur propre conscience active* » (*id.*: 6). Les alternatives à ces études n'ont pas été plus satisfaisantes puisque, bien que prenant les audiences comme des consommateurs actifs de produits médiatiques, ils considéraient pareillement celles-ci en des termes crûment béhavioristes, selon lesquels les audiences auraient des « besoins » plus ou moins « gratifiés » par des consommables médiatiques. En critique à ces nombreuses perspectives et à leurs assomptions communes, une théorie herméneutique a été développée, dans laquelle les produits médiatiques étaient vus comme des « *textes* » qui peuvent être lus et interprétés, de manière grandement comparable aux artefacts littéraires. Ricœur [1981] est connu pour son modèle en triangle dans lequel l'intention de l'auteur, le texte et la réception du lecteur sont associés pour produire une signification textuelle. Dans une perspective plus fonctionnaliste, le modèle du média *qua* texte a pris la forme d'un processus d'encodage/décodage à même d'être investigué empiriquement de part et d'autre. Les études sur la parole radiodiffusée se sont développées en dehors de la critique sur la relation que le modèle d'encodage/décodage établit entre les productions médiatiques et leurs audiences. Elles pointent vers le « *quelque-chose-qui-manque* » (*missing-what*) du projet herméneutique, qui n'accorde que peu d'attention aux

¹ e.g. Bjelic [2007], Jayyusi [2007a] et Dupret & Ferrié [2007 ; 2008] :

<http://www.socialsciences.manchester.ac.uk/disciplines/sociology/about/events/ethnography/journal/>

circonstances de la production médiatique. En considérant la parole radiodiffusée comme un phénomène de plein droit, ces études insistent sur le fait que « *la structure et le contenu des messages [ne peuvent pas être considérés] indépendamment du média interactif dans lequel ils sont générés* » [Heritage et al. 1988, p. 79-80]

Bien que certains ethnométhodologues aient pu se référer au matériau médiatique en tant que texte [e.g. Jayyusi 1988 ; 1991], ce dernier terme n'y est pas investi du même sens que dans la plupart des théories de la communication. En effet, la métaphore du texte sert seulement à désigner un matériau audiovisuel produit et doté de structures d'intelligibilité et de ses éléments procéduraux. S'engager dans l'étude praxéologique de tels « textes » signifie dès lors non seulement porter attention à la dimension sémantique, mais également se focaliser sur les dimensions catégorielle, de séquentielle et contextuelle dans lesquelles toute pratique de lecture du texte est nécessairement enchâssée. Plutôt que de tenter de deviner ce que l'auteur du texte avait en tête au moment de l'écrire et de le publier, les études praxéologiques adoptent tangentiellement l'attitude naturelle du lecteur lorsqu'il est confronté à ce texte médiatique. Parce que les textes sont conçus pour être lus et compris, leur complexité n'implique pas une opacité mais, au contraire, une accessibilité aux lecteurs qui le lisent avec les moyens qu'ils utilisent pour comprendre l'ordre et les propriétés du monde social et naturel [Jayyusi 1984, p. 289]. Dans ce qui suit, nous posons dans un premier temps les fondations d'une analyse ethnométhodologique du vidéoclip. Puis, dans un second temps, nous analysons en détails les vidéoclips sélectionnés et décrivons leur socio-logique aux niveaux du plan, de la séquence et de la narration globale.

1. La socio-logique du vidéoclip

L'on doit à Wittgenstein l'esquisse d'études explorant le travail des jeux de langage dont les grammaires pratiques sont spécifiques à des groupes de gens dont la vie façonne et, en même temps, est façonnée par les activités dans lesquelles ils sont engagés et dans les contextes dans lesquels ils ont lieu. La production médiatique est l'un de ces nombreux jeux de langage dont les particularités, les propriétés, l'organisation et l'intelligibilité peuvent être décrites. Dans un article qui fait date, Lena Jayyusi [1988] a établi les fondements d'une socio-logique de ce qu'elle nomme le « *texte filmique* ». Bien que l'on puisse se demander si la notion de film *qua* texte soit des plus appropriées, ce travail demeure une contribution majeure qui établit le principe de « *l'intelligibilité scénique du monde social* », la logique audiovisuelle des comptes (*accounts*) filmiques, et leur caractère séquentiel. Paul Jalbert [1999] a complété ce tableau des études ethnométhodologiques sur les médias avec la notion de « *structure d'intelligibilité* ». Et l'on doit à Eric Livingston [1995] des études pionnières dans l'anthropologie de la lecture, dans lesquelles il développe la notion d'« *action instruite* » où l'activité consistant à regarder est faite d'une paire (appelée paire *Lebenswelt*) constituée analytiquement, et non phénoménologiquement, des instructions et de la pratique de ces instructions. Dans cette section, nous présentons plus en détails les blocs constitutifs de notre cadre d'analyse.

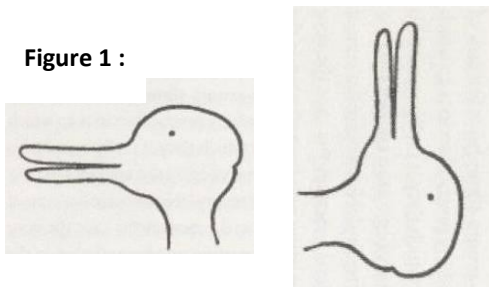
C'est à travers l'examen du langage et de ses mécanismes que l'on peut décrire adéquatement la manière dont les vidéoclips acquièrent leur sens routinier et non problématique². Ces mécanismes constituent une grammaire, dans le sens original et non formel du mot [Wittgenstein 2005 ; McGinn 1997]. Alors que les grammaires formelles, bien qu'originellement établies sur l'arrière-plan de pratiques langagières effectives, prescrivent le « bon usage », la grammaire à laquelle nous nous référons est descriptive et n'a aucune

² Par non-problématique, nous voulons dire pleinement et aisément accessible à une audience moyenne.

posture normative : nous l'appelons une grammaire pratique puisqu'elle reflète la structure de la pratique d'utilisation du langage, c'est-à-dire les schèmes du langage tel qu'employé. Les grammaires pratiques et formelles entrent souvent en conflit, et cela est principalement dû à la tendance de la seconde à abstraire des règles générales à partir de cas singuliers et, dès lors, à effacer les éléments spécifiques et significatifs de configurations singulières. Cependant, ce sont les modes distinctifs de fonctionnement du langage, dans ses nombreux domaines, qui nous montrent la nature du phénomène qu'il vise à décrire. Parmi ces domaines, il y a les différents médias et leurs genres respectifs, entre autre l'audiovisuel, où les images, les sons et les textes se combinent et produisent des formulations qui sont alignées sur les méthodes techniques de la production ordinaire de sens. En d'autres termes, les productions de médias audiovisuels appartiennent à une forme de vie, c'est-à-dire à un ensemble de pratiques au cours desquelles des combinaisons particulières de langage sont spécifiquement développées de manière à produire du sens sur fond d'une structure d'arrière-plan — ou d'un horizon — de compréhension. Ce n'est, en aucune manière, un monde idéal où les expressions suivent des règles formelles, mais un éventail entier de contextes réels dans lesquels le langage joue des fonctions particulières et permet une compréhension commune en fonction des spécificités des contextes pratiques. C'est pourquoi nous devons décrire les mécanismes utilisés pour produire et faire sens de ces accomptes médiatiques.

De manière ethnométhodologiquement respecifiée, nous nous engageons à étudier « *la connaissance et les procédures que les membres sociétaux utilisent pour son assemblage, son test, sa gestion et sa transmission* » [Garfinkel 1990, p. 77], c'est-à-dire la grammaire pratique que ces procédures rendent accomptables et pour lesquelles elles sont accomptables. Pour nos objectifs présents, nous nous concentrons sur la grammaire pratique du fait de voir et de regarder des vidéoclips dans le but de décrire « *les multiples connexions entre des mots et d'autres mots, phrases et expressions tels qu'ils sont utilisés par des 'maîtres du langage naturel', et les multiples connexions entre des types d'expression et des sortes de circonstances pendant lesquelles et à propos desquelles elles peuvent être utilisées* » [Coulter 1989, p. 49]. Cela revient à répudier les approches cognitivistes du « sens », c'est-à-dire à rapatrier ce concept dans le domaine de ses accomplissements pratiques. Cela n'implique pas de tenter de dévoiler ce qui reste « sous la surface » de ces vidéoclips, mais plutôt d'adopter une attitude ordinaire vers les productions médiatiques familières, nous permettant ainsi de voir ce qui est « donné à voir ». Cela tient de l'expérience visuelle dont Wittgenstein nous dit qu'elle consiste à discerner des aspects. Voir des objets est accompli en contexte à travers le fait de remarquer des aspects. On pourrait appeler cela un processus noétique ou de *Gestalt* : toute vision d'un objet est construite à travers la reconnaissance primordiale d'un tout, et à travers l'émergence de ces aspects selon leur relation à ce tout. Si, pour quelque raison que ce soit, cette relation des aspects au tout est sapée, nous faisons face à une crise, quelque chose de l'ordre de la rupture, qui appellera en retour à une forme de réparation (e.g. un accompte sur les causes de la brèche ou l'identification d'un autre type de relation aspects-tout, comme dans le fameux exemple du canard-lapin ; figure 1)

Figure 1 :



À travers l'analyse de grammaires pratiques engagées dans la formulation et la compréhension des vidéoclips, nous conduisons un type d'anthropologie du « voir » et, plus spécifiquement, d'anthropologie du « regarder ». Regarder un programme télévisé n'est pas un processus analytique, mais dans une attitude naturelle de saisie d'un flux d'images et de sons dans leur continuité et dans l'unité qu'ils cherchent à produire [Livingston 1995]. Le téléspectateur ne déconstruit pas chacun des éléments des vidéoclips qu'il regarde, mais il les prend comme un tout dont les éléments variés forment, ensemble et de manière inséparable, sa structure d'intelligibilité³. Toute une variété d'indices scéniques et contextuels converge vers la production de cette texture d'intelligibilité de l'objet télévisé. Ces indices, et leur dépendance mutuelle, ne procèdent pas d'un seul vidéoclip, mais de l'activité consistant à regarder le vidéoclip. Les indices de compréhension des vidéoclips sont d'ores et déjà présents et sont découverts et assemblés par le téléspectateur. En d'autres mots, le travail consistant à regarder un clip revient à chercher l'organisation de ce regard que le vidéoclip décrit. Il ne s'agit donc pas de considérer le clip d'un côté et, de l'autre, le fait de le regarder, dans une correspondance plus ou moins totale, mais bien de s'en saisir comme étant une seule et même chose, à savoir l'action de regarder un vidéoclip, que l'on pourrait concevoir comme une paire unifiant les données filmiques et l'activité consistant à les regarder (la paire clip/lecture du clip). Le clip est toujours enchâssé dans le travail consistant à regarder le clip et, au cœur même de ce travail, il fournit les éléments indiquant comment il doit être compris. Au même moment, le travail consistant à regarder le clip est irrémédiablement enchâssé dans le clip, et il réside dans le travail consistant à découvrir comment le clip prescrit la compréhension qu'il entend produire. À propos de lecture textuelle, Livingston [1995, p. 15] explique : « *De cette manière, un texte procure un 'acompte' de sa propre lecture ; le texte est un 'compte-rendu de lecture', une histoire à propos de comment sa propre lecture devrait être faite* ». Le travail du regard révèle sa nature d'activité dans la transparence, l'intelligence, l'ambiguïté et le non-sens grammatical du vidéoclip et la mesure dans laquelle le clip est regardé de cette manière. Ce travail se manifeste également, en pratique, par le fait que l'action de regarder trouve, dans le clip, les raisons de continuer sur la ligne des premières étapes qui sont toujours organisées dans cet objectif. Regarder le clip est un travail en continu qui requiert des compétences préalables, mais pas de connaissance préalable, et il est continuellement accompli dans la relation essentielle entre le clip et l'action de le regarder par le téléspectateur. Cela opère à la manière d'une recherche par le téléspectateur de l'entreprise inductive de son concepteur. La texture du clip induit une direction dans le regard du téléspectateur, et ce regard active rétrospectivement l'induction opérationnelle dans le clip, et ceci est disponible dans la perspective de l'attitude naturelle et ordinaire de l'action de regarder de tout téléspectateur compétent.

Le téléspectateur compétent n'éprouve absolument aucun problème à identifier le genre auquel ces vidéoclips appartiennent ni, de plus, ce qui les différencie d'autres programmes comme le journal télévisé, les publicités ou les émissions sportives. La distinction qu'il fait résulte de la texture *Gestalt* du travail du regard. En suivant à nouveau Livingston [1995], nous affirmons que chaque clip ne procure pas seulement les éléments sémantiques de sa compréhension, mais aussi les indications contextuelles de son genre et, ainsi, de l'intelligibilité qu'il doit avoir. Continuons à suivre Livingston. La littérature enfantine peut déployer de multiples indicateurs contextuels et n'évite pas la répétition, parce qu'elle s'adresse à un lecteur qui est encore en train de développer ses compétences de lecture. On peut dire dans ce cas que la littérature pour enfants « surdétermine » son interprétation. Un texte technique, pour sa part, utilise également de multiples indicateurs contextuels, mais il

³ Selon Wittgenstein [2005, p. 212] : « *ce que je perçois dans l'émergence d'un aspect n'est pas une propriété de l'objet, mais une relation interne entre cet objet et d'autres objets* ».

évite la répétition ; il semble dès lors « proportionnel » à son interprétation. Un poème, alors qu'il multiplie lui aussi ses indicateurs, aura tendance à les fragmenter, à les éparpiller et à obscurcir la relation qu'ils entretiennent les uns avec les autres. On dira dans ce cas qu'il « sous-détermine » sa lecture. Comme la littérature enfantine, le vidéoclip tend à surdéterminer la lecture qui peut en être faite. Il multiplie les indices contextuels et n'évite pas la répétition. Il propose une forme d'argumentation qui, par l'abondance des éléments pointant dans une direction, conduit à une conclusion surdéterminée. Ceci dit, le clip procure également les éléments qui permettent au téléspectateur de le distinguer (et, par là même, de distinguer son concepteur) des autres qui partagent avec lui un air de famille, c'est-à-dire une ressemblance de genre (pas nécessairement substantielle) dans la famille des vidéoclips de plaidoyer, par exemple. Comme nous le verrons, il est difficile de confondre les deux vidéoclips que nous analysons : l'orientation idéologique de leurs diffuseurs et les audiences qu'ils visent à atteindre sont clairement surdéterminées.

Les vidéoclips que nous analysons partagent une intelligibilité et une accountability naturelles. Comme Metz [1974, p. 145] le souligne justement : « *ce qui doit être compris, c'est que les films sont compris* ». Il s'agit probablement d'un truisme, mais il a le mérite de rendre remarquable le caractère ordinairement vu mais pas remarqué des productions médiatiques. En regardant une séquence filmée depuis la perspective d'une attitude naturelle, celle de la vie de tous les jours, nous la comprenons avec des ressources et des moyens identiques à ceux que nous utilisons pour comprendre l'ordre et les caractéristiques du monde social et naturel [Jayyusi 1988, p. 289]. Regarder la télévision, si nous nous limitons à cet aspect de la sociologie visuelle qui nous concerne dans cet article, ne se fait pas dans un vide, mais de manière contextuelle et située. Par « contextuelle », nous entendons que l'activité n'est pas indépendante des contraintes de la configuration de production et de réception vers laquelle les membres s'orientent ; par « située », nous voulons dire que cela s'appuie sur l'arrière-plan de connaissance partagé par les téléspectateurs, ce que l'on pourrait appeler leur culture. La cohérence des vidéoclips dépend de nos « ressources d'intelligibilité (télé)visuelle », parmi lesquelles notre capacité d'identification, de catégorisation et d'inférence. Le téléspectateur compétent reconnaît le visage du personnel politique le plus en vue, il fait la différence entre des personnes mal intentionnées et des personnes innocentes, et infère la victimisation en associant des femmes, des enfants et des ruines. La cohérence du texte télévisuel « *n'est dès lors pas une cohérence analytique formelle, mais plutôt une organisation d'objectivités pratiques trouvées dans le caractère scéniquement reconnaissable de choses telles que des cours d'action, des relations visibles, des routines familières, etc.* » [Macbeth 1999, p. 148]. Ce n'est donc pas tant le résultat d'opérations codées que celui de scènes et d'activités reconnues et comprises pour ce qu'elles sont par un téléspectateur ordinairement compétent : elles sont réflexivement disponibles dans nos descriptions et nos comptes ; elles sont les caractéristiques reconnaissables de la vie quotidienne telle qu'elle est vécue ; elles sont enchâssées dans nos manières de voir et d'agir. Jayyusi [1988, p. 273] parle à cet égard de la « *transparence scénique* » du monde social et des images que nous avons de lui. Cela ne veut pas dire qu'aucun problème de compréhension ne peut subsister, mais que ces mêmes problèmes forment la partie intégrale du caractère reconnaissable du monde. Ainsi, contrairement à une vision communément acceptée, les images ne possèdent pas un nombre infini de significations. L'action possède sa cohérence visuelle et l'image n'est pas radicalement abstraite du contexte de son visionnage : on peut voir dans la rue l'image d'une fête d'anniversaire et trouver dans son organisation scénique une « proto-narration » de qui ces gens sont, comment la scène en est arrivée à ce qu'elle est et comment elle en est venue à être cela pour nous [Jayyusi 1991, p. 149]. La scène offre un nombre limité d'interprétations possibles que le téléspectateur saisira en ayant recours aux ressources

disponibles à un membre ordinaire de la société habituelle des gens qui peuvent être amenés à regarder ce type de scène.

En considérant « *la structure et les procédés d'intelligibilité hérités du film en même temps que les engagements d'arrière-plan qui interagissent avec certaines de ces structures et procédés, de manière à générer une possible compréhension donnée et une évaluation de celle-ci* » [Jalbert 1999, p. 37], nous nous engageons dans une description de la socio-logique à l'œuvre dans la compréhension d'un matériau vidéo spécifique. En paraphrasant Jalbert [*ibid.*], nous soutenons que, au vu des données vidéo (*V*), leur analyse et la description des engagements d'arrière-plan pertinents (*B*), on peut arriver à une compréhension spécifique (*U*) telle que possiblement générée par *V* sur l'arrière-plan de *B*. Cependant, à la différence de Jalbert [*id.*, p. 39], nous soutenons qu'il est possible, au moins dans le matériau que nous avons sous la main, d'attribuer à leurs diffuseurs particuliers la production délibérée d'une compréhension *U* spécifique des données *V*. Ces diffuseurs, en effet, à travers certaines de leurs productions, mettent en avant de manière accomptable leur identité de diffuseur et s'auto-attribuent des audiences [Dupret, Ferrié 2007]. Puisque ces productions médiatiques sont intentionnellement conçues pour convoier un message spécifique, leur grammaire, c'est-à-dire la combinaison accomptable d'images, de sons, de voix et de textes appuyés sur des engagements d'arrière-plan leur confère un contenu défini. Cependant, pour accéder à un tel contenu, nous devons montrer comment ces productions offrent une structure de possibilités tout en instruisant, au même moment, une seule et unique compréhension. Cela requiert de décrire les nombreux plans, individuellement, séquentiellement et comme un tout, et d'inspecter ce qu'ils montrent, ce qui est inséparable de comment ils montrent ce qu'ils montrent [Macbeth 1999, p. 165].

2. La grammaire pratique de vidéo-clips contrastifs à propos du terrorisme

Il est à présent temps de nous tourner vers notre matériau que nous ne prenons pas comme des échantillons de quoi que ce soit, mais comme de claires instances de production médiatique. Ce type d'analyse est praxéologique dès lors qu'elle implique l'élucidation de structures de sens telles qu'elles sont disponibles et compréhensibles dans l'action de visionner les vidéoclips. Elle concerne la grammaire pratique de ces clips et la socio-logique qui l'organise à plusieurs niveaux, parmi lesquels les catégories sélectionnées, les présuppositions et implications, les inférences pratiques, l'organisation séquentielle, la forme d'énonciation et la narrativisation.

Ce matériau est constitué de deux vidéoclips concernant le thème du terrorisme au Moyen-Orient. Le premier pouvait être vu sur la chaîne d'information *Al-Arabiya*, connue pour être financée par des mécènes saoudiens, alors que le second était disponible sur *Al-Manar*, mieux connu en tant que portevoix du Hezbollah libanais⁴. Les deux chaînes proposent de courtes narrations condamnant le terrorisme, soit de manière explicite à travers le slogan « Le terrorisme n'a pas de religion », soit de manière implicite à travers l'ascription de la qualité de « terroriste » à une catégorie de personnes spécifiquement désignée. Elles prennent la forme de clips d'une minute, souvent présentés dans des périodes d'interlude, combinant des images, du texte, des sons et de la musique, de manière à compléter un message porteur d'une condamnation.

⁴ Bien que l'on pourrait ajouter de nombreux éléments à propos de ces deux chaînes satellites, nous restons fidèles à la perspective de l'attitude naturelle du téléspectateur et renvoyons le lecteur à la page d'introduction que l'on peut trouver (en Anglais) sur le site Internet respectifs de ces chaînes (cf. webographie).

Plans

Puisque cette section est consacrée à des images isolées, il semble pertinent de l'introduire avec quelques remarques à propos des photographies et de leur représentation de scènes à l'intérieur du monde. Doivent-elles être considérées dans les termes d'un paradoxe entre nature (la trace de quelque chose) et culture (une construction sociale) ? Comme L. Jayyusi [1991] l'indique, « *la tendance théorique principale a consisté à mettre l'emphase sur le caractère 'culturel' des objets photographiques envers et contre l'illusion de leur statut 'naturel'* », et cela s'est fait au détriment de l'analyse de l'intelligibilité photographique. Premièrement, en insistant sur l'image en tant qu'artefact culturel, on risque de manquer un élément en filigrane mais évident de la « lecture » photographique, à savoir le fait qu'elle prend racine dans les pratiques sociales et naturelles de raisonnement. Deuxièmement, la nature théoriquement ouverte de l'interprétation de photos s'avère beaucoup moins polysémique quand elles sont prises dans leur dimension praxéologique : dans de nombreux cas, le sens se révèle être instruit et surdéterminé.

Prises isolément, les actions-dans-l'image ont à la fois une immédiate disponibilité au regard et une trajectoire indépendante. Cela signifie qu'elles peuvent être vues pour ce qu'elles sont dans l'image, telles que disponibles à un spectateur ordinairement compétent, et que, cependant, elles sont simultanément enchevêtrées dans une trajectoire qui, de manière compréhensible, a commencé avant que la caméra ne tourne et après qu'elle a cessé de tourner. Dans la Figure 2, capture d'image NT03/47 du vidéoclip *No Terror* (NT, par la suite), nous voyons de jeunes garçons jouer au football : nous le comprenons pour ce que c'est, et nous savons également que leur partie de football a commencé avant que la caméra ne soit en action et qu'elle s'est finie après cela (et ceci n'est pas rendu disponible par la caméra). Malgré le fait que la plus grande partie de la trajectoire des actions-dans-l'image est située en dehors de ce que la caméra rend directement disponible, ce n'est pas seulement l'image qui est comprise, mais également sa trajectoire complète. Cette structure *Gestalt* de la lecture de l'image peut être expliquée par le fait que l'image, dans et à travers elle-même, enclenche des dispositifs de catégorisation qui vont bien au-delà de la disponibilité visuelle [Jayyusi 1991]. Dans notre cas, parmi les nombreux procédés de catégorisation qui sont rendus disponibles, nous pouvons mentionner les suivants : « pratique sportive » (jouer au football), « âge de la vie » (jeunes gens), « genre » (garçons), « géographie locale » (football de rue), « géographie culturelle » (un quartier populaire arabe). Tous s'avèrent ordinairement visibles et accessibles, et, effectivement « *scéniquement transparents* » (*ibid.*). Dans la capture NT19/47, c'est plutôt le seul dispositif de la « nation » (le drapeau irakien) qui est mobilisé, et le fait qu'il soit à terre infère sa chute et induit la menace ou le risque qu'il soit piétiné. Dans la capture NT29/47, nous observons les dispositifs « notabilités » (des dirigeants religieux et civils), qui sont collaborativement appariés, et « figures d'autorité » (le pouvoir par la sagesse ou le pouvoir par les armes), qui sont oppositionnellement appariés. Ici, à nouveau, il y a une trajectoire complète qui est rendue disponible, bien qu'elle ne soit pas directement accessible par l'image : il y avait quelqu'un pour brandir le drapeau avant qu'il ne soit jeté à terre, et il y aura quelqu'un pour le ramasser et le brandir à nouveau, malgré les ennemis et leur propension à chercher à mettre à terre la nation irakienne ; le quartier est composé de strates sociales, incluant des notabilités civiles et religieuses, et leur robustesse, leur sagesse et leur solidarité leur permet de contrer la force brute.

Figure 2 – No Terror: plans isolés

NT03/47 garçons jouant au football



NT19/47 drapeau irakien à terre



NT29/47 notabilités face aux armes



Ceci est également vrai des images du vidéoclip d'*Al-Manar* (MN, par la suite), telles qu'extraites dans la Figure 3. La capture MN03/19 montre le visage d'un bébé dans un cercueil au moment de ses funérailles. Ceci est particulièrement déductif et inductif, en ce sens que la scène qui est capturée indexe nécessairement, d'une part, la violence qui est à l'origine de sa mort et, d'autre part, les manifestations massives qui ont accompagné son enterrement. La capture MN08/19 est l'image d'une foule applaudissant à la destruction d'une mosquée, et ceci se réfère de manière contingente à l'occupation de la Palestine et aux abus commis par l'occupant. Enfin, la capture MN11/19 montre les ruines d'une maison, et elle n'est qu'un moment saisi de la trajectoire s'étendant du bombardement de zones civiles en Palestine à la dépossession d'une famille désormais sans abri.

Figure 2 – Al-Manar: plans isolés

MN03/19 bébé dans un cercueil



MN08/19 la foule



MN11/19 maison en ruine



Indépendamment de la place qu'elles occuperont dans les séquences spécifiques et les narrations globales, ces images véhiculent leur propre intelligibilité. La capture NT03/47, tout en montrant des garçons jouant au football, est auto-suffisante pour dire au téléspectateur que la scène est celle de la vie quotidienne de gens ordinaires dans quelque environnement urbain irakien. Nous ne savons pas ce qui est arrivé à la partie de football avant et après son moment filmé et nous n'avons pas besoin de le savoir pour reconnaître et comprendre la scène. En ce sens, nous pouvons parler d'images isolées en tant que narrations autonomes. Ces « narrations à-plan-unique » possèdent un caractère allusif permettant de les caractériser en tant que proto-narration ou — mieux, puisque ceci n'est pas seulement lié à l'ordonnement temporel — narrations allusives et sous-déterminées. Ainsi, l'intelligibilité de l'image est-elle donnée non seulement au regard des particularités immédiatement disponibles à l'œil mais également des particularité « hors-caméra » qui ne sont pas visibles « *mais cependant praxéologiquement présentes* » : « *La trajectoire naturelle et effective de l'action est invisible, bien qu'elle 'anime' et qu'elle soit incorporée à l'intérieur de ce qui est visible dans ce cadre, et bien qu'elle soit lue dans les termes de l'histoire naturelle des particularités qui sont visibles* » (*ibid.*). Il y a une narration de la partie de football qui est rendue disponible, bien qu'elle ne soit pas directement visible ; et il y a une narration des maisons de civils dans MN11/47, bien

que nous ne puissions voir que les blocs de béton restant après le bombardement. Il y a également une narration de la normalité du lieu, de la scène et des gens, à laquelle il est fait allusion dans l'image, bien que cela ne soit jamais dit en autant de mots : la normalité d'une partie de football entre garçons dans une rue populaire (NT03/47) ou la normalité des vêtements des gens (NT29/47) ; ou même, la normalité de la mort (MN03/47) et de la destruction (MN11/19) dans un environnement palestinien, bien que, dans ce cas-ci, la normalité se construise comme une non-normalité consolidée (ce qui devrait être le cas en des temps exceptionnels n'est qu'une scène normale dans ce même contexte). Il est à noter que les trajectoires prévisibles de la narration normale d'un objet peuvent induire en erreur. Ainsi en va-t-il du drapeau (NT19/47) : dans ce cas, l'objet est cadré dans une situation qui peut être considérée comme anormale au regard des valeurs spécifiques et, partant, de l'activité normale liée à celui-ci.

La familiarité joue un rôle particulier dans le processus consistant à rendre praxéologiquement présents les trajectoires et les cours d'action qui ne sont pas explicitement présents dans le cadre de la caméra. Les lieux, les visages, les scènes, les codes vestimentaires n'indexent pas seulement des catégories de personnes ou de lieux, mais également des activités qui sont liées à ces catégories. Dans les images de *No Terror* (Figure 2), nous voyons comment les scènes (NT03/47 : une partie ordinaire de football de rue entre garçons dans une rue populaire), les objets (NT19/47 : le drapeau irakien à terre), les codes vestimentaires et les visages stéréotypiques (NT29/47 : des notabilités dans leur costume traditionnel), posent le cadre de la narration d'une manière qui indexe un environnement irakien ordinaire bien que stéréotypique et les nombreuses activités d'une journée ordinaire dans une rue ordinaire. Dans les images d'*Al-Manar* (Figure 3), la familiarité est d'un autre type, du fait du recours à des archives documentaires. C'est une familiarité avec des situations réelles et qu'on peut avoir vues dans un environnement palestinien : l'enterrement d'un bébé (MN03/19), une foule en train de crier (MN08/19) et les ruines d'une maison (MN11/19). Ces images saisissent un moment spécifique de trajectoires d'action rendues praxéologiquement présentes mais matériellement invisibles : le bombardement d'innocents, la profanation de lieux saints, la destruction de maison de civils. Ces images indexent également un lieu, la Palestine, bien que de manière radicalement différente de ce qui est fait avec *No Terror* : c'est l'indexation d'une histoire, c'est-à-dire que cela vise des événements dont l'histoire ne peut être contestée (comme ceci est spécifiquement indiqué par la référence de ces images en tant qu'images non-fictionnelles ou « de terrain », à travers la mention en haut à gauche du mot « direct », « *mubâshir* » en arabe). En d'autres termes, alors que les images *No Terror* induisent le contexte d'un quartier populaire irakien en mettant en scène ses éléments stéréotypiques, les images d'*al-Manar* déduisent un contexte palestinien à partir du reportage, au sens premier du terme, d'événements historiques-*qua*-stéréotypiques.

Les photographies ne sont pas lues dans un vide : « *Il y a toujours et inévitablement un contexte qui fournit et rend disponible, de manière éventuellement implicite, un accompte du caractère observable et reportable de la 'scène'* » (*ibid.*). Les légendes, pendant ou à la fin de chaque vidéoclip, jouent souvent ce rôle de désambiguïsation : elles sont réflexivement accomptables de la lisibilité/observabilité/reportabilité de l'image. La chaîne de télévision que le spectateur regarde procure également un contexte au regard de laquelle le clip est accomptable et qui est lui-même accomptable de la lisibilité du clip : « *C'est ce contexte qui vous fournit ce que j'ai sous le regard à l'intérieur des photos qui sont réimprimées ici* » (*ibid.*). Cependant, ce n'est pas le contexte seul et isolé qui procure cette lisibilité, mais à la fois le photographe et le contexte dans lequel le photographe est rendu visible : leur intelligibilité respective est prédicative l'une de l'autre. En regardant le vidéoclip d'*Al-Manar*, il m'est rappelé (e.g. par le logo de la chaîne en haut à droite de ces images) que je suis en

train de regarder la chaîne *Al-Manar*, c'est-à-dire, une chaîne arabophone dont l'engagement dans la résistance contre Israël est communément connu par à peu près tout le monde et rendu visible de part en part de la programmation de la chaîne. En regardant le vidéoclip sur *Al-Arabiya*, je sais que je suis en train de regarder une chaîne arabophone d'information qui diffuse souvent des spots de promotion de la reconstruction de l'Irak. Autrement dit, la sélection de la chaîne que je regarde est un élément contraignant de mon visionnage du vidéoclip : elle procure le contexte de l'intelligibilité de ce clip alors que, au même moment, le clip lui-même participe de la catégorisation de la chaîne.

Séquences

Chaque plan est en soi une narration allusive. Cependant, dans le vidéoclip, ceci est également disposé de manière séquentielle, dans la juxtaposition avec d'autres plans. En ce sens, c'est une partie d'une séquence émergente en train de se déployer, qui est à la fois distincte et constitutive de la narration globale. Cela révèle à quel point le produit dans son intégralité est un objet laminé, avec les nombreuses trajectoires implicites et explicites de plans, de séquences et de narration globale, entremêlés à travers l'organisation grammaticale complexe des images.

Figure 4 – NoTerror: la séquence des hommes-armés-surgissant-sur-la-place



La Figure 4 illustre l'une des séquences du vidéoclip NoTerror. Elle combine sept plans qui, s'ils sont pris isolément, ont un sens allusif bien qu'incomplet, ce que Hart (1961) appelle dans le champ de la loi, une texture ouverte. La juxtaposition de ces plans instruit la manière avec laquelle chaque trajectoire doit être lue de façon à arriver à un sens cohérent. Autrement dit, alors que l'ensemble des plans nous procure un fond commun de trajectoires indépendantes, leur ordonnancement séquentiel les articule dans une seule et même lecture.

On peut voir un convoi de voitures roulant à vive allure sur une route poussiéreuse (NT04/47). Des hommes armés sont assis à l'arrière d'un pick-up (NT05/47) et, par vertu du fait d'être placé après le plan précédent, ceci est l'image des hommes-armés-assis-à-l'arrière-d'un-pick-up-dans-un-convoi-de-voitures-roulant-à-vive-allure. Autrement dit, la juxtaposition indexe la voiture à l'arrière de laquelle ils sont assis en tant qu'elle est la voiture qui était précédemment montrée dans le plan précédent. En ce sens, la juxtaposition a un effet transitif. Puis, NT06/47 montre le visage d'une fille qui est facilement reconnaissable en tant que visage effrayé. Isolément, ceci poserait la question de la source causale d'une telle peur. Dans une séquence, l'attribution de la peur à la vue d'hommes-armés-assis-à-l'arrière-d'un-pick-up-dans-un-convoi-de-voitures-roulant-à-vive-allure se fait de manière naturelle.

NT05/47 désambiguïse NT06/47, alors que NT06/47 canalise la trajectoire de NT05/47 en instanciant l'une des conséquences possibles de la vue d'hommes armés. Autrement dit, la juxtaposition a un effet causal qui s'accomplit d'une manière interprétative documentaire (l'image-dans-la-séquence prend son sens depuis la narration émergente et elle donne réflexivement à celle-ci son sens). De retour aux voitures en convoi dans NT07/47, nous n'avons aucun problème à les assimiler à ce que nous avons vu dans NT04/47, non pas tant parce que les voitures se ressemblent, mais principalement parce que nous sommes d'ores et déjà immergés dans une trajectoire que nous accompagnons spontanément, au lieu de chercher des alternatives faiblement plausibles. Cela illustre l'une des maximes du lecteur selon Sacks [1996] : si deux images placées en juxtaposition peuvent être lues en tant que la deuxième suit la première, alors lisez-les de la sorte. NT08-10/47 duplique la peur instanciée en NT06/47. Cependant, un autre facteur contribue à la production de cet effet : la bande-son (des tirs d'armes automatiques), qui donne quelques nouvelles explications aux attitudes des gens. Dans ce cas, la causation est produite par la juxtaposition de la vidéo et de l'audio, ce dernier combinant des sons « *in vivo* » (diégétiques) et une musique dramatique. Cela montre que, d'une part, une séquence est organisée de manière laminée, dans laquelle l'une des strates (e.g. l'audio) peut produire causalement l'autre (e.g. la vidéo) ; et, d'autre part, que la narration dans son intégralité est dramatiquement produite afin d'atteindre progressivement un apex prévu bien qu'inconnu.

Figure 5 – Al-Manar: ils ont profané les [lieux] sacrés



La Figure 5 illustre la seconde séquence du vidéoclip d'*Al-Manar*. Elle combine quatre plans qui, s'ils sont pris isolément, ont un sens direct, documentaire et indexical. En fait, ils se réfèrent à des événements historiques : la visite provocative d'Ariel Sharon sur l'esplanade d'al-Aqsa à Jérusalem et la destruction d'une mosquée par une foule de colons dans la Palestine occupée. Prise isolément, chaque image a une trajectoire historique communément connue. Prises dans une séquence, ces images instruisent d'une manière très directe la lecture de la conclusion écrite : "ils ont profané les lieux sacrés".

La première image (MN06/19) identifie le sujet de la séquence en tant qu'il est l'ennemi archétypique. Cette image est composée de questions du type « où, quand, qui, pourquoi et quoi » (*wh-questions*), lesquelles questions pointent conjointement vers une conclusion entrevue mais toujours non spécifiée. Pour n'importe quelle personne familière du Moyen-Orient et de ses conflits, cette image indexe le visage de l'ancien Premier ministre d'Israël, Ariel Sharon (le « qui »), pendant sa célèbre parade (le « quoi ») sur l'esplanade d'al-Aqsa (le « où ») en septembre 2000 (le « quand »), laquelle est perçue par les Arabes comme ayant provoqué à dessein (le « pourquoi ») le second soulèvement palestinien. Cependant, le plan n'est se suffit pas à lui-même, dès lors que rien n'indexe directement le lieu et le moment de

ce qu'il capture. Il repose dès lors sur la compétence du spectateur à l'encadrer dans un horizon de familiarité, de reconnaissabilité et de compréhension. Cet horizon est, dans ce cas, la trajectoire historique bien que cachée (parce qu'elle n'est pas à l'écran) de ce plan, c'est-à-dire ce qui est anticipé et ce qui suit le moment précis de sa capture. Cette image ouvre une série d'inférences possibles quant à ce que Sharon faisait : il paradait, il profanait, il provoquait, etc. Le plan suivant (MN07/19) renforce le sentiment que cela a à voir avec l'atteinte au caractère sacré des lieux de l'islam, alors qu'il montre quelqu'un (qui pourrait contextuellement être vu comme un colon juif), sur le toit de ce qui apparaît être une mosquée (à cause du dôme et du minaret qu'il montre en partie), en train de desceller un parpaing qu'il jette ensuite à terre, face à une foule de juifs israéliens (certains portant la kippa, d'autres brandissant le drapeau israélien), en train d'applaudir à ce spectacle (MN08/19). Bien que ces deux derniers plans indexent un événement historique qui n'a pas la dimension iconique du précédent, ils se réfèrent néanmoins à une série d'éléments itératifs dans l'histoire de la Palestine [cf. Jayyusi, 2007b], rendant cette scène de juifs-se-réjouissant-à-la-vue-de-la-destruction-d'une-mosquée éligible au statut de preuve de ce qu'affiche la légende de la séquence : « ils ont profané les [lieux] sacrés ». La légende fonctionne comme une formulation, c'est-à-dire, quelque chose de semblable à un résumé, une conclusion, une paraphrase du point principal de la parole précédente, qui la préserve, l'efface et la transforme [Heritage & Watson 1979]. La figure de Sharon et la scène du destructeur acquièrent le sens exclusif de profanateurs, annulant, pour les finalités pratique de la démonstration syllogistique acquise par le spot, toute autre interprétation possible. En ce sens, la signification d'une séquence est prospectivement produite par chaque plan et rétrospectivement déterminée par la formulation, laquelle sert de pierre angulaire à l'argument de la narration globale.

Narrations

Si les plans véhiculent un sens propre qui est enchevêtré dans leur organisation morceau par morceau et séquentielle, la somme de ces séquences est productive d'une narration globale. Cette idée de narration globale doit être comprise dans les termes d'une production *Gestalt* tout/parties, c'est-à-dire, le déploiement d'une histoire dont certains détails gagnent en saillance, selon leur signification dans la compréhension de l'ensemble, alors que d'autres détails sont confinés à l'arrière-plan et disparaissent de fait de l'histoire racontée au spectateur. Dans le dessin du lapin-canard (cf. *supra*), nous avons vu comment l'incise prend la valeur singulière d'une bouche, quand elle est placée en bas sur fond d'une tête typique de lapin, alors qu'elle est écartée lorsqu'elle est placée sur la droite sur fond d'une tête typique de canard. De façon similaire, les détails des nombreux plans et séquences prennent une pertinence particulière à cause de la place qu'ils occupent et de la fonction qu'ils acquièrent dans la narration globale, alors que d'autres détails ne recouvrent aucune saillance et sont dès lors généralement écartés.

La Figure 6 tente de détailler les nombreuses étapes à travers lesquelles le vidéoclip *No Terror* produit une narration qui pourrait être résumée de la manière suivante : « Contre la violence terroriste, seuls la fermeté, le courage et la solidarité communautaire peuvent restaurer la nation et faire reculer l'ennemi ». Il est à noter qu'il peut y avoir des sélections alternatives des plans et des séquences qui, ensemble, mènent à la production d'une seule et unique narration. Une narration est une structure en forme de trame, d'une ouverture à une fermeture, à travers les nombreuses étapes d'une histoire présentant une surface immédiatement accessible à n'importe quel lecteur/spectateur compétent. Dans ce cas, cela commence avec une vue surplombante d'un carrefour dans un quartier populaire où des voitures et des gens s'engagent dans leurs affaires quotidiennes (NT01-03/47). Soudain, des voitures font irruption sur la place (NT04-07/47) et effrayent les gens (NT06/47, 08/47). Des

hommes en armes descendent de ces voitures (NT08/47) et commencent à tirer à l'aveuglette, à frapper des personnes âgées et des jeunes gens avec leurs fusils (NT11-23/47). Non moins soudainement, quelque chose se produit et laisse les hommes armés manifestement perplexes (NT24/47). Il y a un homme au milieu de la route, qui se tient seul et fait face aux terroristes alors que les autres continuent de fuir (NT26-28/47). Ferme face à la menace des fusils, cette figure tribale arabe est ensuite rejointe par d'autres notabilités qui se tiennent épaule contre épaule (NT29-30/47) et qui sont eux-mêmes rejoints par d'autres personnes (NT31-35/47), parmi lesquelles un garçon brandissant le drapeau irakien (NT36-38/47). Ensemble, ils font face aux terroristes qui commencent à battre en retraite (NT40-44/47), avec la légende : « Connais ton ennemi ». La narration finit sur l'image surplombante, qui passe au noir et blanc, des terroristes cernés par la foule. Sur cette image, la conclusion surimposée se lit comme suit : « le terrorisme n'a aucune religion, le terrorisme n'a aucune nation » (NT45-47/47).

Figure 6 – NoTerror: le récit « nation *versus* terrorisme »



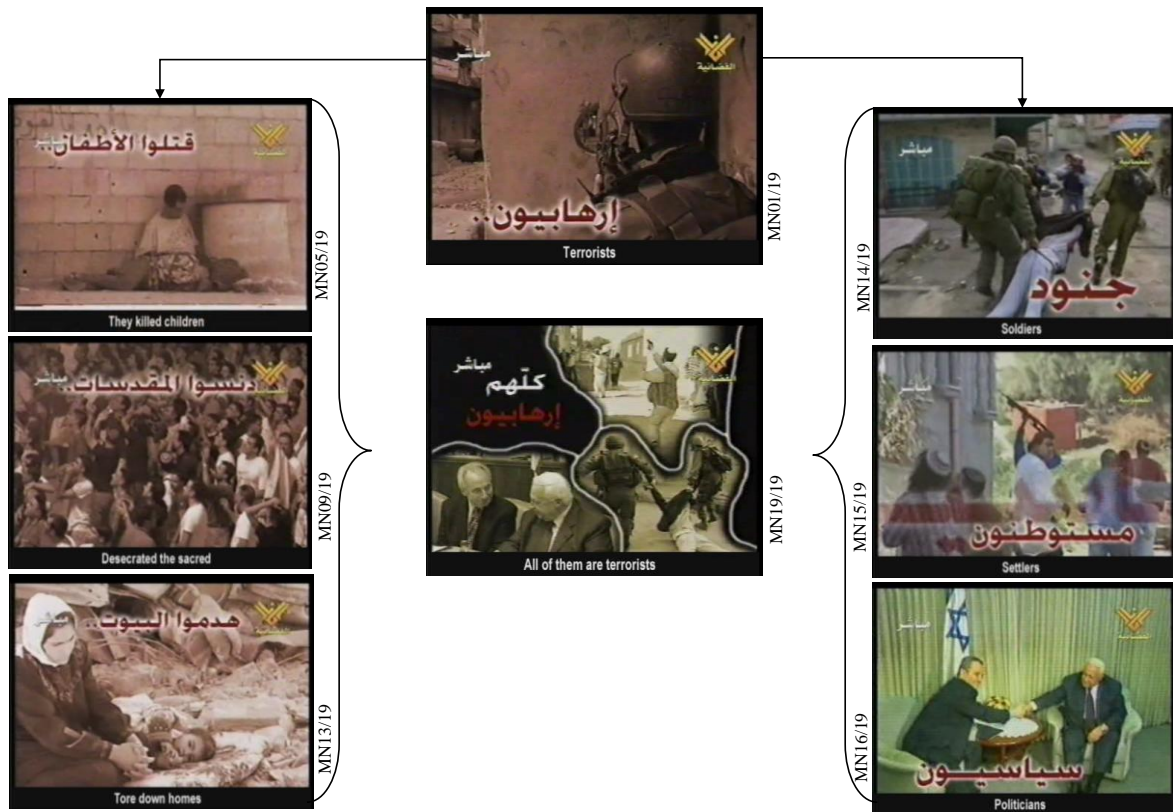
Le vidéoclip est construit comme une narration fictionnelle qui est contrainte par son format d'une minute. Autrement dit, elle montre intentionnellement, et dans un cadre contraint, les protagonistes d'une intrigue dont la force persuasive réside dans sa capacité à présenter une histoire courte et simple, bien que manichéenne. Cette narration est scénarisée et intentionnée par le producteur du spot qui a pris ses plans de manière consistante, c'est-à-dire dans le but pratique et spécifique de produire une histoire courte de plaidoyer. Il y a une sorte de réalisme dans cette production, en ce sens que le film est tourné en extérieur, dans un « vrai » quartier arabe/irakien, dans l'environnement d'une rue dont les particularités semblent « naturelles ». Cependant, l'histoire elle-même ne vise pas spécialement le réalisme : c'est une fable morale. Le producteur accentue les éléments caractéristiques de chacune des nombreuses parties et ne cherche pas à donner un compte-rendu fidèle de la

complexité des figures psychologiques et des situations ; au contraire, il contraste des personnages, des comportements et des situations, de manière à arriver à une conclusion qui contient explicitement un enseignement moral : le courage, la solidarité et l'unité l'emportent nécessairement sur la lâcheté, le factionnalisme et l'esprit de clocher.

La Figure 7 tente de détailler la narration globale du vidéoclip *Al-Manar* qui présente une structure totalement différente de celle de *No Terror*. Elle s'ouvre sur un titre montrant un soldat de dos, à hauteur de sa tête, avec pour légende : « terroristes.. » (MN01/19). Puis, il se poursuit avec deux listes de trois items, eux-mêmes subdivisés en listes de trois plans. Il est à noter que ces listes sont organisées en trois parties. Les listes de trois sont une des techniques de communication les plus efficaces et les plus largement utilisées [Atkinson 1984] ; pour citer Matoesian [2001, p. 92] : « les listes sont des techniques expansives pour produire des airs de famille, créer une unité conceptuelle entre des éléments qui seraient sinon divers dans le champ perceptuel, et pour classer des actions à travers une litanie rythmiquement texturée et progressivement étendue d'items similaires ». Généralement composées de trois items, l'efficacité de ces listes réside dans la production d'un sens d'unité et de complétude et dans le pouvoir exponentiel de leur rythme poétique. Au même moment, elles évitent le danger de la verbosité et conduisent naturellement à une conclusion qui s'instancie comme le vrai dénouement d'un micro-drame construit par la répétition morphologique et sémantique. Dans le vidéoclip *Al-Manar*, il y a deux listes principales de trois formulations visuelles (reconnaisables par la couleur sépia de l'image) et textuelle (qui se lit comme l'ascription d'actes mauvais à un « ils » anonyme) : la première est faite d'activités et la seconde de catégories de personnes. La liste d'activités est elle-même divisée en une autre liste de trois items visuels qui fonctionne comme suit. Premièrement, la succession d'un bébé en sang sur un lit d'hôpital (MN02/19), d'un autre bébé exposé dans son linceul pendant ses funérailles (MN03/19) et d'un garçon cherchant refuge derrière un homme qui appelle à l'aide (MN04/19) amène à la formulation « ils ont tué les enfants (*qatalû al-atfâl*) (MN05/19)⁵. Deuxièmement, la succession de la visite de Sharon (cf. *supra*; MN06/19), de la destruction d'une mosquée (cf. *supra*; MN07/19) et d'une foule se rejouissant face à ce spectacle (cf. *supra*; MN08/19) amène à la formulation « ils ont profané les [lieux] sacrés » (*dannasû al-muqaddasât*) (MN09/19). Troisièmement, la succession d'un bulldozer détruisant un mur de parpaings (MN10/19), des ruines d'un immeuble détruit (MN11/19) et d'une femme et un enfant assis devant ces ruines (MN12/19) amène à la formulation « ils ont détruit les maisons » (*dammarû al-buyût*) (MN13/19). Quant à la liste de catégories de personnes, elle fonctionne de manière moins aléatoire puisque qu'elle va directement à la formulation et au texte incrusté sur les plans. Premièrement, des soldats en uniforme en train de traîner un homme à terre, avec la légende « soldats » (*junûd*) (MN14/19). Deuxièmement, des personnes portant la kippa et armés de fusils se dispersant autour d'un immeuble avec la légende « colons » (*mustawtinûn*) (MN15/19). Troisièmement, Ariel Sharon et Ehud Barak se serrant la main devant le drapeau israélien et le même Sharon avec ses collègues pendant un conseil des ministres, avec la légende « politiciens » (*siyâsiyyûn*) (MN16-17/19). Le plan final est fait d'une image divisée en quatre parties apparaissant les unes à la suite des autres, et dans lesquelles s'insèrent successivement les scènes figées d'un colon armé et portant une kippa (à droite en haut de l'image), de deux soldats traînant un homme à terre (à droite en bas de l'image), d'Ariel Sharon et de Shimon Perez au conseil des ministres (à gauche en bas de l'image) et, enfin, la formulation conclusive (à droite en haut de l'image) stipulant : « Ce sont tous des terroristes » (*kulluhum irhâbiyyûn*) (MN18-19/19).

⁵ Il est à noter que la troisième image a une valeur iconique, puisqu'elle est immédiatement reconnaissable par un spectateur compétent en tant qu'elle montre la mort de Muhammad al-Durra au début de la seconde Intifada.

Figure 7 – Al-Manar: Le récit « ce-sont-tous-des-terroristes »



Le vidéoclip *Al-Manar* a une structure qui s'avère totalement différente de celle de *No Terror*, tant au niveau de l'argumentation que de la documentation. Elle est basée sur des plans extraits d'archives d'information, qui ont déjà servi de supports aux programmes d'information et qui n'ont en aucun cas été pris dans le but du vidéoclip lui-même. Ceci décuple leur force probante, puisqu'ils ne peuvent être écartés comme de simples mises en scène mais qu'au contraire, ils prennent racine dans des événements historiquement fondés et couverts médiatiquement en temps réel. Ce faisant, les images du spot donnent l'impression de rendre compte du réel, au lieu de créer une fiction. Elles reposent sur la valeur de véridiction des images documentaires. Le vidéoclip *Al-Manar* se présente également comme un procédé « énigme-solution » (*puzzle-solution device*). À propos de la question des applaudissements lors de meetings politiques, Heritage et Greatbatch analysent les nombreuses manières par lesquelles les orateurs façonnent leur discours afin de générer l'alignement de l'audience :

... l'orateur commence par établir une sorte d'énigme ou de problème dans la tête des auditeurs ; ensuite, peu après, il propose comme solution à l'énigme une déclaration qui se présente comme le noyau du message qu'il veut faire passer. Le message adoubé est accentué par l'énigme, qui invite l'audience à anticiper ou deviner sa solution et, du même coup, à écouter attentivement la solution de l'orateur quand il la fait connaître. De plus, comme la délivrance de la solution coïncide naturellement avec l'exposé du message politique, l'audience est normalement capable d'anticiper le moment auquel les applaudissements devraient débiter idéalement. Les orateurs aident normalement les audiences à cet effet en donnant à la solution l'aspect d'une phrase déclarative simple et à a forme active [Heritage, Greatbach 1986, p. 127]

Bien que cela ne soit pas généré dans un environnement interactionnel dans lequel un orateur s'adresse directement à une audience dont il cherche l'approbation orale, la structure des arguments est exactement la même dans le cas qui nous occupe : en combinant le mot

« terroristes.. » avec l'image d'un soldat d'une armée régulière, le producteur pose une énigme (qui sont-ils, ces terroristes ?) dans l'esprit des spectateurs. Puis, petit à petit, il présente les éléments qui procurent ensemble la solution. Ces éléments sont organisés de manière syllogistique : (1) les terroristes commettent des actes de méchanceté (ils tuent, profanent et détruisent) ; (2) de tels actes sont commis par ces catégories de personnes (des soldats, des colons, des politiciens) ; (3) ces catégories de personnes sont des terroristes (« Ce sont tous des terroristes »). À l'évidence, le producteur ne cherche pas à s'attirer des applaudissements. Cependant, la structure (e.g. en trois parties) et le caractère simple du message invite également à un alignement (bien que non pragmatique) de la part de l'audience. C'est un message qui induit une sorte de solidarité négative [Ferrié, 2004], c'est-à-dire l'alignement du producteur comme de l'audience sur le type de narration qu'ils attendent les uns des autres et qui, une fois mis en avant, ne peut être mis en concurrence.

Il faut enfin souligner que les narrations globales fonctionnent également comme contexte réflexif des images. Autrement dit, le sens fonctionne de deux manières. Premièrement, c'est un phénomène émergent et séquentiellement organisé, dans lequel un plan prend son sens parce qu'il est positionné après les plans et les séquences précédents, et dans lequel l'intégralité de la narration tire son sens d'un processus chronologique et incrémental. Deuxièmement, et plus important, les particularités du vidéoclip prennent leur signification du sens général rétrospectivement produit par la narration en tant que telle. Dans l'éventail d'interprétations et de trajectoires possibles pour chaque plan, il y en a une et seulement une qui permet la narration globale *ex post facto*. Par exemple, un regard peut être décrit comme effrayé parce qu'il est juxtaposé après un plan de voitures roulant à vive allure (NT04-05/47), mais c'est seulement parce que nous connaissons finalement toute l'histoire que, de fait, ce regard peut être perçu comme un regard effrayé, dont la cause se trouve dans les voitures — et non, par exemple, le tir d'un ballon avec lequel les enfants sont en train de jouer. Le *Gestalt* de la narration procure le contexte contraignant de la compréhension des plans et des séquences du clip.

Conclusion

Cet article s'attachait à la description des narrations de vidéoclips en tant que ceux-ci constituent des jeux de langage. Ces vidéoclips visent à catégoriser le monde de la vie sur la question du terrorisme, de manière à démontrer la position du diffuseur vis-à-vis de cette question et à induire le point de vue de son audience. Le jeu de langage du vidéoclip de plaidoyer se caractérise par le déploiement d'une grammaire pratique, c'est-à-dire l'agencement d'images, de sons et de textes de manière complexe bien que structurée, dans une structure sous forme de trame contrainte par la spécificité de l'activité en cours autant qu'orientée vers elle. Plusieurs éléments contribuent à la formation de cette grammaire : des éléments techniques spécifiques à la communication audiovisuelle ; des éléments de format spécifiques au genre du vidéoclip (tenir un argument idéologique en moins d'une minute) ; des éléments structurels spécifiques au type narratif choisi (fictionnel ou documentaire). Cette grammaire est la condition même de l'intelligibilité partagée des narrations pour leurs producteurs et leurs audiences. En d'autres termes, l'audibilité de ces messages de plaidoyer est conditionnée par une telle grammaire pratique.

Bibliographie

ATKINSON Max [1984] *Our Masters' Voices. The Language and Body Language of Politics*, London, Methuen

BJELIC Dušan [2007], “Edited Identities and geopolitics of Global Media”, *Ethnographic Studies*, 9

COULTER Jeff [1989] *Mind in Action*, Cambridge/Oxford, Polity Press/Basil Blackwell

DUPRET Baudouin, FERRIE Jean-Noël [2007] “The Audience They Assign Themselves : Three Arab Channels and Their ‘Self-Presentation’ (al-Jazeera, al-Manar, al-Hurra)”, *Ethnographic Studies*, 9

— [2008] “News Headlines: Stating in Brief what is relevant in Today’s World (al-Arabiyya, al-Jazeera, al-Manar, BBC World)”, *Ethnographic Studies*, 10

FERRIE Jean-Noël [2004] *Le régime de civilité en Égypte. Public et réislamisation*, Paris, CNRS Éditions

GARFINKEL Harold [1990] *Studies in Ethnomethodology*, 1ère edition 1967, Cambridge/Oxford, Polity Press/Basil Blackwell

HART Herbert [1961] *The Concept of Law*, Oxford, Oxford University Press

HERITAGE John, CLAYMAN Steven, ZIMMERMAN Don [1988] “Discourse and Message Analysis” in HAWKINS R., WIEMANN J., PINGREE S. (dirs.) *Advancing Communication Science: Merging Mass and Interpersonal Processes*, Sage Annual Reviews of Communication Research, 16

HERITAGE John, GREATBATCH David [1986] “Generating Applause: A Study of Rhetoric and Response at Party Political Conferences”, *The American Journal of Sociology*, 92 (1)

HERITAGE John, WATSON Rod (1979) “Formulations as Conversational Objects” in PSATHAS G. (dir.), *Everyday Language*, New York, Irvington Press

HUTCHBY Ian (2006) *Media Talk: Conversation Analysis and the Study of Broadcasting*, Maidenhead, Open University Press

JALBERT Paul [1999] *Media Studies: Ethnomethodological Approaches*, Lanham/New York/Oxford, University of America

JAYYUSI Lena [1984] *Categorization and the Moral Order*, Boston, Routledge and Kegan Paul

— [1988] “The equivocal Text and the Objective World: An Ethnomethodological Analysis of a News Report”, *Australian Journal of Media and Culture*, 5 (1). Disponible en ligne : cf. webographie

— [1991] “The Reflexive Nexus: Photo-practice and Natural History”, *The Australian Journal of Media and Culture*, 6 (2). Disponible en ligne : cf. webographie

— [2007a] “Between saying and Showing: Making and Contesting Truth Claims in the Media”, *Ethnographic Studies*, 9

—— [2007b] “Iterability, Cumulativity, and Presence: The relational Figures of Palestinian Memory” SA’DI A. & ABU-LUGHOD L. (dirs.) *Nakba. Palestine, 1948, and the Claims of Memory*, New York, Columbia University Press

LIVINGSTON Eric [1995] *An Anthropology of Reading*, Bloomington, Indiana University Press

MACBETH Douglas [1999] “Glances, trances and Their relevance for a Visual Sociology” in JALBERT Paul (dir.) *Media Studies: Ethnomethodological Approaches*, Lanham/New York/Oxford, University of America

McGINN Marie [1997] *Wittgenstein and the Philosophical Investigations*, London and New York, Routledge

METZ Christian [1974] *Language and Cinema*, La Haye, Mouton

MATOESIAN Gregory [2001] *Law and the Language of Identity: Discourse in the William Kennedy Smith Rape Trial*, Oxford, Oxford University Press

RICOEUR Paul [1981] *Temps et récit*, Paris, Seuil

SACKS Harvey [1996] *Lectures in Conversation*, Oxford, Basil Blackwell

WITTGENSTEIN Ludwig [2005] *Recherches philosophiques*, 1^{ère} édition 1953, Paris, Gallimard, « Bibliothèque de philosophie »

Webographie

http://www.alarabiya.net/en/about_aa_tv.html

<http://www.almanar.com.lb/NewsSite/AboutUs.aspx?language=en>

<http://wwwmcc.murdoch.edu.au/ReadingRoom/5.1/Jayyusi.html>

<http://wwwmcc.murdoch.edu.au/ReadingRoom/6.2/Jayyusi.html>

<http://www.noterror.info/OurCampaign.aspx?ID=274&MID=30&PID=0>

<http://www.socialsciences.manchester.ac.uk/disciplines/sociology/about/events/ethnography/journal/>