



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE DELLA RICERCA

Alma Mater Studiorum Università di Bologna Archivio istituzionale della ricerca

La relazione tra suono e immagine nel videoclip

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Lucio Spaziante (2022). La relazione tra suono e immagine nel videoclip. *ESTETICA. STUDI E RICERCHE*, 2, 313-330 [10.14648/105532].

Availability:

This version is available at: <https://hdl.handle.net/11585/919706> since: 2023-03-01

Published:

DOI: <http://doi.org/10.14648/105532>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

Lucio Spaziante

La relazione tra suono e immagine nel videoclip

Abstract

The article is dedicated to the analysis, even with semiotic tools, of music video as a media format, specifically in its textual features, particularly in the relationship between sound and image, also highlighting gaps still existing in the observation of this matter. Music video belongs to the more general audiovisual field, but has its own specificity due to the relevance possessed by the song, and therefore by the music, with respect to film footage. This demonstrates how is overturned the usual hierarchy which in audiovisual language usually gives a primary role to images. The article takes up some elements of the aesthetic debate on the role of music in cinema, and then explores aspects relating to the conventionality of the synchronization between song and video, showing its concrete separateness, through a description of the diegetic sound logics, also presenting analyzes of two case studies.

Keywords

Music video, audiovisuals, sound, image, text.

L'articolo nasce dal tentativo di approfondire alcuni aspetti del linguaggio audiovisivo nel videoclip, cercando di comprendere alcuni degli elementi specifici di questo formato mediale, e quali siano le possibili ricadute sulle teorie che si sono interrogate, più in generale, sulla relazione tra suono e immagine in movimento.

Prima di entrare nel merito di questioni legate al linguaggio e alla testualità, a scopo introduttivo è utile ricordare che, al di là di antenati storici che si sono già serviti della relazione tra musica e formati cinematografici (come Cinebox o Scopitone¹, oppure video promozionali realizzati dai Beatles), il videoclip si afferma realmente, come prodotto industriale di natura comunicativa attorno alla metà degli anni Ottanta del Novecento, allo scopo di ampliare la diffusione di un brano musicale, e come mezzo promozionale dell'identità di un'artista o di una band in un contesto televisivo. Infatti, in uno dei primi compendi di studi sul tema, si parla di "music television", sottolineando il ruolo del *medium* piuttosto che il *formato* mediale² nello specifico.

La golden age è legata all'esperienza televisiva di MTV, quando sulla copertina di *Time* del 26 dicembre 1983 campeggiava la scritta "Video rocks! A musical revolution", e il videoclip veniva individuato come la tendenza culturale del momento. Il suo rapido sviluppo avverrà secondo stringenti logiche commerciali, contribuendo ad una enorme produzione quantitativa, che darà luogo anche a percorsi creativi e talvolta

¹ B. Di Marino, *Segni sogni suoni: quarant'anni di videoclip da David Bowie a Lady Gaga*, Mimesis, Milano 2018, p. 17.

² S. Frith, A. Goodwin, L. Grossberg, L. (Eds.), *Sound and vision: The music video reader*, Routledge, London 1993, p. ix; G. Sibilla, *Dancing in the Dark. MTV, Music Videos, Bruce Springsteen and the Aesthetics of Rock in the Eighties*, in «ZoneModa Journal», 9(1), 2019, pp. 107–118.

sperimentali³. Il videoclip diventerà altresì un laboratorio sul linguaggio audiovisivo, i cui effetti si potranno cogliere soprattutto nei decenni successivi.

Lo studio del fenomeno ha vissuto un picco di interesse nel periodo tra la fine degli Novanta e l'inizio degli anni Zero⁴, sia come fenomeno musicale e mediale, sia per la sua valorizzazione come forma artistica. Poi, con il declino, dovuto sia motivi commerciali e industriali, sia per una generale perdita di interesse, la letteratura sul tema si è significativamente ridotta, per poi riprendere in questi ultimi anni⁵. Oggi il videoclip da un punto di vista culturale non è più un fenomeno di tendenza: ma se da una parte rappresenta un capitolo della storia dei media e della *popular music*, dall'altra, essendo migrato dal broadcast televisivo verso Internet, molta della fruizione musicale contemporanea passa ancora attraverso videoclip, su YouTube o più di recente su TikTok, con numeri di visualizzazione tra i più alti in assoluto⁶. Dunque non si tratta solo di un tema utile alla storia dei media, bensì di un fenomeno nel pieno del mainstream, sebbene con un ruolo differente rispetto al passato.

Come sottolinea Will Straw in un articolo che fa il punto sull'argomento, a trent'anni dalle sue prime riflessioni, il videoclip è divenuto attualmente

both ubiquitous and minor. Through the videoclip, YouTube has become the largest disseminator of music in the world, but this has diminished rather than magnified the clip's status as cultural form.⁷

La formula testuale del videoclip

La formula testuale del videoclip prevede che, partendo dalla canzone – ovvero da un testo sonoro “preesistente”⁸ – vi si associ successivamente un materiale visivo, che andrà a costituire l'effettivo testo finale, composto assieme dalla canzone e dalle immagini⁹. Questo formato definisce a priori una posizione gerarchicamente superiore

³ M. Goddard, *Audiovision and Gesamtkunstwerk: The Aesthetics of First and Second Generation Industrial Music Video*, in G. Arnold, D. Cookney, K. Fairclough, and M. Goddard, M. (Eds.) *Music/Video: Histories, Aesthetics, Media*, Bloomsbury Academic, New York 2017, pp.163-180.

⁴ G. Sibilla, *I linguaggi della musica pop*, Bompiani, Milano 2003; C. Vernallis, *Experiencing Music Video: Esthetics and Cultural Context*, Columbia University Press, New York 2004; A. Amaducci, S. Arcagni, *Music video*, Kaplan, Torino 2007; H. Keazor, T. Wübbena (Eds.), *Rewind, Play, Fast Forward: The Past, Present and Future of the Music Video*, Transcript Verlag, Bielefeld 2010.

⁵ L. A. Burns, S. Hawkins (Eds.), *The Bloomsbury Handbook of Popular Music Video Analysis*, Bloomsbury Academic, New York 2019; B. Osborn, *Interpreting Music Video: Popular Music in the Post-MTV Era*, Routledge, New York 2021.

⁶ T. Jirsa, M.B. Korsgaard, *The Music Video in Transformation: Notes on A Hybrid Audiovisual Configuration*, in «Music, Sound, and the Moving Image», 13(2), 2019, 111-122, p. 112.

⁷ W. Straw, *Music Video in its Contexts: 30 Years Later*, in «Volume !» online, 14-2-2018, mis en ligne le 01 janvier 2021, consultato il 10 giugno 2022. URL: <http://journals.openedition.org/volume/5696>.

⁸ P. Peverini, *Il videoclip*, Meltemi, Roma 2004, p. 73.

⁹ Per approfondire le dinamiche tra immagini e video nel videoclip, soprattutto in chiave narrativa, si veda anche L. Marconi, *Enunciazioni e azioni nella canzone e nel videoclip*, in «E|C» Serie Speciale, Anno I, n. 1, 2007, pp. 105-113.

della canzone, e dunque dell'espressione sonora, rispetto all'immagine in movimento¹⁰. La canzone di norma non viene modificata, ed è esclusivamente a partire da essa che verrà successivamente creato il materiale visivo.

Dunque, trattandosi di un testo *costruito su un altro testo*, esso presenta una consustanziale natura *derivativa* e al contempo *traduttiva*.¹¹ L'ideazione e la realizzazione visiva (e poi narrativa) vengono infatti create in modo multiforme *a partire dalla* canzone, ovvero da quell'insieme eterogeneo in essa contenuto, composto da musiche, voci, parole (*lyrics*), e che nel suo insieme costituisce il materiale di ispirazione: un potenziale bagaglio a partire dal quale costruire significati articolati e sincretici¹².

Si pone dunque una questione sulla natura della correlazione tra la dimensione sonora e quella visiva: le immagini definiscono possibili *riferimenti* alla canzone, o comunque entrano in una relazione con essa. Ma il panorama dei videoclip si presenta, del resto, estremamente variegato, con le più diverse tipologie di *connessione* tra canzone e video, che vanno dall'associazione arbitraria e immotivata (ovvero immagini quasi "casuali"), fino alle più sofisticate strategie di interrelazione testuale tra i due piani.¹³

La realizzazione di un videoclip talvolta conserva dunque un carattere a metà tra l'artigianale e l'"usa e getta", procedendo quasi in modo improvvisato, attraverso la semplice "giustapposizione" delle immagini alla musica¹⁴.

Il primato della canzone rispetto alle immagini ha definito in ogni caso una sorta di "contratto libero", grazie al quale la stretta correlazione tra immagini e suoni non risulta così necessaria. E già nella sua fase "televisiva" primaria, e ancor più da quando di fatto è stato collocato al di fuori della tv lineare generalista, il videoclip ha spesso rappresentato un formato poco controllato da logiche editoriali o da censure morali, realizzato con bassi budget (specie in determinati periodi) e senza significativi ritorni economici, consentendo le scelte creative più varie, a seconda dei gusti e del posizionamento degli autori delle canzoni.

Da un punto di vista della sua struttura, il videoclip può prevedere anche innesti inediti (non contenuti nella canzone se ascoltata nel formato audio), che possono precedere o seguire la canzone vera e propria, con l'inserimento di dialoghi, effetti sonori, o addirittura musiche aggiuntive, venendosi a costituire talvolta come una sorta di mini-film. Una tendenza sviluppatasi specie con l'avvento di YouTube e dei formati digitali¹⁵. La canzone è il testo-*primario*, contenente già un materiale sonoro eterogeneo (voci, suoni, musiche); a questo si associa un testo-*secondario* - il girato video - che sovrappone alla canzone altro materiale eterogeneo, fatto di immagini ma anche di suoni; il tutto andrà a comporre il videoclip in quanto testo finale stratificato. Un testo

¹⁰ A. Gaudin, *Le clip comme forme d'expression musico-visuelle: pour une esthétique de la relation musique-images*, in «Volume !», 14 (2), 2018, pp. 97-110.

¹¹ N. Dusi, L. Spaziante (a cura), *Remix-Remake: pratiche di replicabilità*, Meltemi, Roma 2006.

¹² C. Vernallis, *Experiencing Music Video: Aesthetics and Cultural Context*, Columbia University Press, New York 2004, p. 13.

¹³ P. Peverini, *Il videoclip*, Meltemi, Roma, 2004, p. 62.

¹⁴ E. Caston, *'The First Cut is the Deepest' Excerpts from a Focus Group on Editing Music Videos, with Explanatory Historical and Theoretical Notes*, in «Music, Sound, and the Moving Image», 11(1), 2007, 99-118, p. 112.

¹⁵ *Id.*

che dunque solo in apparenza può essere considerato semplice, o banale, visto che deriva da un processo creativo e produttivo complesso, anche sul piano del linguaggio.

Quella stretta gerarchia che pone la canzone a monte del video, potrebbe far ipotizzare che le possibilità espressive del medium siano limitate, ma la realtà è ben diversa, se non addirittura opposta. Il videoclip presenta un'ampia libertà sul piano del linguaggio visivo, proprio grazie al fatto che è la canzone a fornire una garanzia di *continuità*: il sonoro procede, infatti, senza interruzioni, qualsiasi cosa si veda. La materia sonora, mantenendo una propria omogeneità semantica, conferisce elementi di *coerenza* ma anche di *coesione*, perché porta avanti stile e andamento iniziali, sino alla fine del testo. Dato che il compito di preservare la stabilità strutturale viene assolto dalla canzone, la dimensione visiva risulterà libera di organizzarsi secondo regole molto più libere, rispetto - ad esempio - a quelle del cinema.

Viene così *invertita* la tradizionale gerarchia dell'audiovisivo cinematografico, per la quale è l'immagine in movimento a prevalere sulla dimensione sonora¹⁶. Nel rapporto tra visivo e sonoro, infatti, il cinema ha storicamente optato per una *stretta correlazione* tra i due piani, dove il suono possiede prevalentemente la funzione subordinata di fornire elementi di "realismo" a ciò che è visibile sullo schermo¹⁷. Oltre agli effetti sonori, il cinema naturalmente si serve anche della musica in qualità di *colonna sonora*, ma anche qui conferendole una relazione di sudditanza rispetto all'immagine¹⁸.

Nel videoclip, inoltre, la relazione tra audio e video non è rigidamente connessa tramite la *sincronizzazione* dei dialoghi e dei rumori, come è tipico del cinema. La sincronizzazione è sì contemplata, in particolare quella labiale connessa al canto (*lip sync*), ma non costituisce un vincolo effettivo¹⁹.

È come se, al suo apparire nella parte finale del Novecento, il videoclip avesse compiuto un salto temporale all'indietro, al momento della firma del contratto convenzionale tra audio e video nel cinema, e avesse offerto un'alternativa possibile a questo modello. Questo è senza dubbio uno dei motivi di successo del videoclip come medium artistico, e della sua stessa influenza sul cinema, in quanto territorio di sperimentazione sul linguaggio.

Lo spazio sonoro nel videoclip

Lo spazio sonoro del videoclip – con il quale intendo idealmente le "fonti sonore" diegetiche o extra-diegetiche dalle quali vengono attinti i suoni che possiamo ascoltare - è occupato per la totalità, o quasi, dalla canzone. Dunque di norma non è contemplato l'ascolto del parlato, e della recitazione, salvo nei prologhi o negli epiloghi aggiuntivi. Come dicevamo, la canzone possiede uno statuto apparentemente privilegiato, perché è su di essa che si costruisce il video. Ma se alla canzone non è previsto che si apportino modifiche, né che si aggiungano suoni, ciò implica che nel videoclip la dimensione sonora di fatto risulti una superficie *chiusa e inviolabile*, mentre quella visiva sia invece

¹⁶ M.B. Korsgaard, *Music Video after MTV: Audiovisual Studies, New Media, and Popular Music*, Routledge, London 2017, p. 48.

¹⁷ M. Chion, *L'audio-vision: son et image au cinéma*, Armand Colin, Paris 1991.

¹⁸ J. Buhler, *Theories of the Soundtrack*, Oxford University Press, New York 2018.

¹⁹ P. Peverini, *Il videoclip*, Meltemi, Roma 2004, p. 74.

tutta da inventare. La chiusura agli interventi sonori impedisce, o comunque riduce al minimo, ogni forma di realismo sonoro. La narrazione sonora nel cinema, infatti, poggia essenzialmente sul supporto degli effetti e dei rumori (ad esempio i passi di un personaggio, o la chiusura della portiera di una macchina), nella loro sincronizzazione e correlazione alle immagini. Nel caso dei videoclip, invece, è assente la presenza dei suoni *diegetici* (ovvero quelli interni all'universo narrato), tranne rare eccezioni. Il piano visivo si va dunque a costituire come una successione di immagini che possono o meno essere semanticamente correlate al piano sonoro (suoni o parole cantate), ma senza una condivisione fenomenica dei due universi tra loro: i suoni “della canzone” non si incontrano con i suoni “del video”. Ovvero, l'universo diegetico sonoro del video non comunica affatto, o quasi per nulla, con l'universo diegetico sonoro della canzone. Pensiamo ad esempio al rumore dei passi: se in un film è normale udire i passi di un personaggio che cammina, in un videoclip, durante la riproduzione della canzone, il rumore di quei passi di norma *non* si sente. I due universi (la canzone e le immagini) vengono riprodotti in modo associato, ma senza che si compenetrino l'un l'altro. Per cercare di spiegare ulteriormente questo concetto, semplice ma a suo modo non immediato, è come se, guardando una partita di calcio, abbassassimo il volume della telecronaca televisiva, e la sostituissimo con il commento radiofonico dello stesso evento: i commenti della radio e le immagini della tv sarebbero, in questo caso, due universi sonori paralleli ma tra loro non comunicanti.

Si aggiunga anche che nell'azione scenica del videoclip, non necessariamente i personaggi del video “ascoltano” la musica che noi spettatori sentiamo. Costoro compiono azioni che solamente noi siamo in grado di osservare accompagnate da una canzone, che è collocata in una sorta di universo parallelo. E siamo nuovamente solo noi spettatori, in uno stato di eterna onniscienza, ad avere il privilegio testuale di accedere a entrambi.

Ciò spiega in parte perché il mondo visivo del videoclip segua spesso logiche non realistiche, talvolta oniriche, in ogni caso poco legate a un principio di causalità. Proprio perché viene di continuo creata un'associazione tra due universi (quello sonoro e quello visivo) che può essere del tutto arbitraria, o comunque non collegata, e dunque poco logica, e poco realistica. La comunicazione tra i due universi diegetici avviene per il tramite di una *meta-enunciazione* di cui è garante e partecipe solo lo spettatore. Cioè, siamo noi spettatori che, tramite la fruizione del videoclip nella sua totalità, accettiamo di associare due universi testuali, musica e immagini, che in realtà sono del tutto *separati*.

Tramite queste riflessioni, diventano più evidenti gli elementi di interesse del videoclip in quanto formato mediale, in quanto da un lato ci dimostra la natura composita e “innaturale” del linguaggio audiovisivo, e dall'altro ne dimostra le molteplici possibilità espressive, derivanti dalla complessa relazione tra suono e immagine. L'audiovisivo conferma ancora una volta di possedere le caratteristiche di un linguaggio, ovvero la capacità di produrre senso, sebbene con caratteristiche non assimilabili a quelle di un linguaggio verbale.

Rapporto suono-immagine: accordo o separazione?

Sulla relazione tra suono e immagine esiste un'ampia letteratura, in particolare relativa all'ambito cinematografico²⁰. La questione trae origine sin dall'epoca del passaggio dal film muto al film sonoro, essendo il medium cinematografico basato fino a quel momento esclusivamente sulla sequenza e sul montaggio di immagini. Quando ad esse venne associata una banda sonora sulla quale poter inserire voci, musiche ed effetti sonori, l'immagine mantenne comunque un ruolo dominante, a cui secondariamente si associava il suono. Nel videoclip, come abbiamo visto, la prospettiva si inverte, e si può dire che si tratti di uno dei rari casi in cui nell'audiovisivo avvenga questo ribaltamento. È per questo motivo che, analizzando la relazione tra suono e immagine nel videoclip, il dibattito cinematografico risulta di estremo interesse.

L'audiovisivo è generalmente basato su una convenzione linguistica per la quale due entità separate – immagini in movimento e suoni – vengono presentate come “naturalmente” *unificate*, quando di fatto non lo sono. Semplicemente, esse vengono *fruite* in modo associato: il girato filmico e la registrazione del suono vengono riprodotti assieme, producendo così l'effetto di un'esperienza unificata²¹. Se il cinema, da parte sua, tendenzialmente *nasconde* la natura autonoma delle due materie espressive, nel videoclip invece, come abbiamo visto, questa condizione separata risulta molto più evidente.

La prassi cinematografica classica si è fondata sulla *sincronizzazione* tra immagine e suono, ovvero su una sorta di effetto mimetico di naturalità simulata, che nel tempo è stata poi assunta come del tutto “normale”. Già a partire dalle prime riflessioni sulla questione, ad esempio quella di Eisenstein²², questa relazione mimetica era però stata messa in discussione, originando una opposizione teorica tra *parallelismo* e *contrappunto*, ovvero tra: modalità in cui il carattere dell'immagine e quello del suono apparivano *convergenti*; e altre nelle quali essi *divergevano* in modo dissonante. Il realismo generato, appunto, dalla visione convergente, si manifestava come un effetto costruito, ed è per questo che divenne oggetto di critiche da parte di chi, invece, vedeva nel medium cinematografico uno strumento critico e di svelamento artistico, piuttosto che conciliante e ridondante.

La sincronizzazione tra immagine e suono è divenuta una modalità discorsiva convenzionale alla quale M. Chion ha dato il nome di *sincretismo audiovisivo*²³. Ovvero, una sorta di contratto “di ferro” che conferisce grande efficacia e potere espressivo alla sincronia. Un meccanismo che contiene elementi di *riferimento* (tramite i suoni

²⁰ C. Gorbman, *Unheard melodies: narrative film music*, BFI Publishing London; Indiana University Press, Bloomington 1987; R. Calabretto, *Appunti per una possibile teoria della musica per film*, in «Musica e storia», 17(3), 2009, pp. 511-566.

²¹ J. Buhler, *Theories of the Soundtrack*, Oxford University Press, New York 2018, p. 12.

²² P. Moormann, *Music and sound as counterpoint? Aesthetic discussions about film music in Germany triggered by the advent of sound film in the late 1920s*, in «Musica e storia», 17(3), 2009, pp. 599-606.

²³ M. Chion, *L'audio-vision: son et image au cinéma*, Armand Colin, Paris 1991.

“indessicali”²⁴, ovvero esplicitamente “riferiti” e connessi alle immagini sullo schermo) e al contempo di *simulazione*, (è altrettanto evidente che l’associazione tra immagini e suoni sia il frutto di una costruzione tecnica), dunque con aspetti di una più vasta problematica semiotica. Da un lato si fa credere che il suono che ascoltiamo sia *collegato* all’immagine che vediamo, o addirittura, che il suono “derivi” dall’immagine stessa. Dall’altro è possibile realizzare una strategia opposta, altrettanto rilevante semioticamente, ovvero quella di *svelare* l’artificio linguistico, connettendo immagini e suoni che *non* hanno *alcuna* evidente relazione tra loro. Ad esempio, associare l’immagine di uno spazio vuoto al rumoreggiare di una folla, oppure - spostandosi sul piano della rottura delle convenzioni culturali - corredare una scena violenta con una musica allegra ed energetica (si pensi ad *Arancia meccanica* di S. Kubrick, USA-UK 1971, oppure a *Le iene* di Q. Tarantino, USA 1992).

Le soluzioni adottate nel rapporto tra suono e immagine hanno originato nel tempo considerazioni estetiche tra loro molto diverse. La sincronizzazione, che nel cinema si è affermata come la modalità standard, per Eisenstein, Pudovkin e Alexandrov ²⁵ ha “distrutto la cultura del montaggio”. A parere degli autori, le immagini mute, montate in modo coeso tra loro, possedevano una loro intrinseca potenzialità narrativa che viene invece annullata dall’ancoraggio con il suono.

Questa osservazione, apparentemente “antica”, risulta invece attuale nel momento in cui la adattiamo al contesto del videoclip, nel quale, come già detto, accade molto spesso che il parlato, ovvero il cantato, *non* venga sincronizzato con l’immagine. Di conseguenza al *montaggio*, che nel videoclip è solitamente costruito ritmicamente sulla canzone (“on the beat”)²⁶, è lasciato un margine di significazione relativamente indipendente. In questo modo si viene a realizzare almeno in parte la prospettiva di Eisenstein, secondo la quale l’unico uso appropriato del suono doveva seguire la logica del *contrappunto*, ovvero quella di una radicale scissione con l’immagine.

Tra le varie prospettive teoriche che si sono susseguite in questa materia, in quella di H. Eisler e T. Adorno²⁷ emerge invece, al contrario, una prospettiva radicale che criticava anche la logica del contrappunto di Eisenstein. Eisler e Adorno sottolineavano che, pur ipotizzando una rigorosa discordanza tra immagine e suono, in qualche misura veniva ugualmente postulata l’esistenza di una *relazione morfologica* tra suono e immagine. Contrari per principio ad un uso ridondante della musica nei film, Eisler e Adorno concepivano invece una sorta di irresolubile differenza tra il dominio dell’immagine e quello del suono.

Al di là dell’accettabilità o meno di questa ipotesi, risulta in ogni caso interessante l’aver stimolato una domanda ulteriore: associare suono e immagine, implica che si produca una sorta di *dimensione terza* che chiamiamo audiovisivo? In altre parole, in

²⁴ J. Buhler, *Theories of the Soundtrack*, Oxford University Press, New York 2018, pp. 12 -15; L. Spaziante, «Sound, image and fake realism», in L. Elleström, O. Fischer, C. Ljungberg (Eds.), *Iconic investigations*, John Benjamins Publishing Company, Amsterdam Philadelphia, 2013, pp. 263 – 274.

²⁵ S. Eisenstein, V. Pudovkin, G. Alexandrov, «A Statement On Sound», in S. Eisenstein, *Selected Works vol. 1 Writings, 1922-34*, I.B. Tauris, London-New York 2010, pp. 113-114.

²⁶ C. Vernallis, *Experiencing Music Video: Esthetics and Cultural Context*, Columbia University Press, New York 2004, p. 27.

²⁷ T. W. Adorno, H. Eisler, *Composing for the Films*, Continuum, London 2007.

che misura è possibile ipotizzare una sorta di reciproca relazione di *risemantizzazione* tra i due domini, tale da generare un risultato univoco? È ciò che in qualche misura è emerso nel dibattito sul rapporto tra musica immagine, influenzato dalla diffusione dei media audiovisivi. La fruizione di tv, spot pubblicitari, videoclip - e ancor di più con l'affermarsi dei media digitali - ha favorito una visione "organica" del linguaggio audiovisivo, dove ogni elemento – visivo e sonoro nei suoi singoli componenti - contribuisce assieme agli altri a determinare un risultato, e dunque un significato, unitario²⁸.

Logiche della relazione tra musica e video

Analizzando la relazione tra musica e immagine nel linguaggio del videoclip, ci sono alcune logiche che in letteratura appaiono convergere tra differenti autori, sebbene con sfumature diverse. Si presentano come categorie "a maglie larghe", che non scendono cioè nel dettaglio delle strategie discorsive. D'altra parte, però, hanno avuto lo scopo di creare un primo ordine di ragionamento attorno all'oggetto.

Andrew Goodwin individua, ad esempio, tre categorie per le diverse relazioni tra suono e immagine nei video musicali²⁹:

1. le immagini possono fungere da *illustrazione* della musica, senza aggiungere nulla di completamente nuovo ai "significati" già presenti nella musica.
2. possono fungere da *amplificazione* della musica, aggiungendo nuovi elementi al significato complessivo senza disordinare il significato musicale.
3. le immagini possono segnare una *disgiunzione*, lavorando in modo dirompente in opposizione alla musica.

In modo simile Nicholas Cook³⁰ definisce la relazione tra suono e immagine come basata su *conformità*, *complemento* o *opposizione*³¹. Anche David Machin lavora con un insieme di relazioni in qualche modo corrispondente, in cui la relazione è di *elaborazione*, *estensione* o *miglioramento*³². Dunque approssimativamente si possono visualizzare queste tre posizioni che in buona parte accomunano i tre autori.

Goodwin	<i>illustrazione</i>	<i>amplificazione</i>	<i>disgiunzione</i>
Cook	<i>conformità</i>	<i>complemento</i>	<i>opposizione</i>

²⁸ N. Cook, *Analysing musical multimedia*, Oxford University Press, New York 1998.

²⁹ A. Goodwin, *Dancing in the Distraction Factory: Music Television and Popular Culture*, University of Minnesota Press, Minneapolis 1992.

³⁰ N. Cook, *Analysing musical multimedia*, Oxford University Press, New York, 1998.

³¹ M.B. Korsgaard, *Music Video after MTV: Audiovisual Studies, New Media, and Popular Music*, Routledge, London 2017, p. 92.

³² D. Machin, *Analysing Popular Music: Image, Sound, Text*, SAGE, London and Los Angeles 2010.

Machin	<i>elaborazione</i>	<i>estensione miglioramento</i>	
--------	---------------------	-------------------------------------	--

La prima categoria è dunque quella dell'*illustrazione*, basata su una relazione di tipo descrittivo: nel caso del videoclip indicherebbe il modo attraverso il quale le immagini sarebbero in grado di *descrivere* la musica. Nel caso del cinema, invece, il rapporto era identico, ma all'inverso: Eisler e Adorno³³ nel 1947 citavano la frase “birdie sings, music sings”, all'epoca adoperata a Hollywood: se nelle immagini del film l'uccellino canta, allora anche la musica del film deve cantare. La musica doveva descrivere le immagini attraverso cliché che, nel tempo, tramite convenzioni connotative e di genere, hanno creato associazioni tra determinati timbri o soluzioni melodico-armoniche e altrettanti stati passionali.

Nel caso del videoclip, all'inverso, la modalità dell'illustrazione diventa: ciò che nella musica si “dice”, nelle immagini si deve “vedere”. Può essere il caso delle parole delle canzoni (*lyrics*) e dunque dei temi narrati, oppure delle “atmosfera” passionali, ovvero quegli elementi sensibili sonori che devono prendere corpo nel video.

Nelle forme primigenie di videoclip, ancora ingenue e artigianali, i video risultavano spesso proprio come un'illustrazione *figurativa* delle parole della canzone, secondo una logica di pura *conformità*. C'è un celebre caso, realizzato da un maestro indiscusso del cinema come Michelangelo Antonioni, per la canzone *Fotoromanza* (1984) di Gianna Nannini, divenuto nel tempo l'esempio di ciò che il videoclip *non* doveva essere³⁴. Ad ogni figura evocata dalle parole dalla canzone (“camera a gas”, “palazzo che brucia”) corrispondeva una resa figurativa letterale sul piano visivo.

Ma già all'interno di un innovativo programma televisivo quale era *Odeon* (RAI, 1976-1978), che conteneva ampi spazi dedicati alla musica, si poteva vedere una sorta di videoclip ante-litteram per *Bomba o non bomba* di Antonello Venditti³⁵. Qui, in corrispondenza del verso “a Sasso Marconi incontrammo una ragazza”, si vedeva il cartello autostradale di Sasso Marconi, seguito poi da una ragazza, seduta su una panchina³⁶. E il resto del video seguiva la medesima logica.

Il videoclip può dunque *illustrare* o sviluppare ciò che viene “detto” nelle parole della canzone, così come può invece tentare di costruire una resa visiva dell'espressione musicale.

Ma se una strategia illustrativa troppo letterale risulta pleonastica, è possibile allora lavorare secondo un'altra logica: quella dell'*amplificazione*, ovvero dell'*elaborazione-estensione* “a partire da”. Elementi verbali o sonori della canzone, dai quali cercare di produrre un'espressione visiva conforme, ma differente. Questa è la logica più diffusa, che cioè vede le immagini come un completamento, o una variazione rispetto a quanto

³³ T. W. Adorno, H. Eisler, *Composing for the Films*, Continuum, London 2007.

³⁴ B. Di Marino, *Segni sogni suoni: quarant'anni di videoclip da David Bowie a Lady Gaga*, Mimesis, Milano 2018, p. 39.

³⁵ *Ivi*, p. 270.

³⁶ <https://www.raiplay.it/video/2022/05/Antonello---Blob---Puntata-del-04062022-21d1a187-ebc2-4617-8c1a-1e0594542bf8.html>

la canzone già contiene. Naturalmente, all'interno di questa categoria si possono mettere in atto le più differenti strategie di connessione tra suono e immagine. Per poi a giungere invece all'estremo opposto, ovvero alle strategie di progressiva, o totale, *divergenza* tra suono e immagine.

Il caso *Star Guitar*: traduzione suono-immagine

Nel saggio di M. B. Korsgaard³⁷, che allo stato attuale rappresenta una delle trattazioni teoriche più complete nel campo del videoclip, si cita il caso di *Star Guitar* (2002) dei Chemical Brothers. Un esempio che vale la pena riprendere, in quanto si tratta di un caso di virtuosismo audiovisivo, che dimostra efficacemente le potenzialità del formato videoclip.

La rilevanza di *Star Guitar* sta nel mostrare un piano visivo che è progettato per realizzare una sorta di corrispondenza punto a punto delle modulazioni e delle variazioni che si apprezzano sul piano sonoro. Una sorta di meticolosa *visualizzazione* del suono, portata avanti attraverso una strategia di *traduzione del suono in immagine*, creativa e al tempo stesso ironica il cui livello qualitativo non sorprende, trattandosi di Michel Gondry, tra i più noti e apprezzati registi di videoclip, nonché del film di culto *Eternal Sunshine of a Spotless Mind* (USA, 2004).

Andando nel dettaglio, il brano dei Chemical Brothers appartiene al macro-genere della musica elettronica dance, non presentando uno sviluppo particolarmente approfondito del testo verbale. Si ripetono frasi come “dovresti provare ciò che io provo, dovresti vedere ciò che io vedo, dovresti prendere ciò che prendo”, il che rende superflua, di fatto, ogni necessità interpretativa delle *lyrics*, salvo un generico appello alla condivisione percettiva. Sarà dunque l'*espressione musicale*, in senso stretto, l'unico elemento coinvolto nel processo di traduzione verso la dimensione visiva.

L'ambientazione del video consiste in una continua visualizzazione del panorama visibile dal finestrino di un treno, sulle linee ferroviarie francesi tra Nîmes e Valence, via Avignone, in diversi momenti della giornata. L'andamento musicale del brano possiede in effetti una configurazione che può evocare una somiglianza ritmica con quella del rumore di un treno, contribuendo ad una sorta di effetto mimetico.

Ma la caratteristica originale del video, con tratti di vera genialità, è quella di avere operato una resa - sincronizzata sul piano visivo - della variazione degli eventi che si ascoltano sul piano sonoro. Come molti brani di musica elettronica, *Star Guitar* si sviluppa per *accumulazione* progressiva di voci strumentali, e poi con improvvise *rarefazioni* o interruzioni; oppure tramite per *variazioni* timbriche, o mediante accrescimento e diminuzione di *intensità* nel volume sonoro. Il tutto, seguendo una logica ritmica e continua, simile appunto alla configurazione acustica del rumore ritmico di un treno: caratteristiche, come si vede, tipicamente sonore, non traducibili in modo immediato sul piano visivo.

³⁷ M.B. Korsgaard, *Music Video after MTV: Audiovisual Studies, New Media, and Popular Music*, Routledge, London 2017.

La soluzione creativa di Gondry è stata allora quella di operare una correlazione diretta tra parametri sonori ed elementi visivi: ad ogni voce strumentale (rullante, cassa, sequencer, hi-hat, ecc.), viene associata una diversa figura iconica nel video, secondo un preciso andamento metronomico (soluzione da lui in parte già adottata nel video *Around The World* dei Daft Punk, nel 1997). Oltre a questo impianto di base, avvengono di continuo variazioni nel paesaggio, che si ripetono anch'esse ritmicamente. Ad esempio, secondo una sorta di codice interno, ad ogni colpo di *clap* appare un palo della luce, oppure ad ogni colpo di cassa appare una cisterna. Oppure, se aumenta il numero di voci strumentali, aumenta la vastità del paesaggio (ad esempio ai due binari improvvisamente se ne associano altri); o ancora, alle variazioni timbriche di un filtro di *phasing*, corrisponde figurativamente l'innalzamento di un muro, oppure il passaggio dal giorno alla notte. Oppure, in corrispondenza dell'ingresso di *voci umane*, il paesaggio diventa quello di una stazione ferroviaria popolata di *persone*. Molte sequenze visive vengono, inoltre, ripetute in modo identico, emulando la configurazione ripetitiva delle sequenze musicali³⁸.

Nell'esprimere visivamente questa tecnica compositiva - l'aggiunta continua di ulteriori tracce - le immagini del video funzionano come una sorta di notazione grafica, non del tutto dissimile dalle interfacce visive conosciute dai software musicali basati su griglia, dove toni, ritmi e altri elementi sono tracciati su una griglia visiva che rappresenta spazialmente il loro posto temporale nella musica.

In quanto tale, le immagini del video sono sorprendentemente parallele al design di molti software di produzione musicale, in cui la musica viene presentata come una serie di tracce orizzontali che corrono in parallelo sullo schermo³⁹.

Il valore esemplificativo di *Star Guitar* di Gondry risiede nel dimostrare le potenzialità del linguaggio del videoclip, attraverso una scelta di forte conformità tra i due piani, evitando però ogni ridondanza descrittiva. Si potrebbe ipotizzare che l'efficacia in questo caso sia derivante da una scelta basata sull'astrazione e sulla *figuralità*⁴⁰. I punti di contatto tra musica e immagine sono basati su una correlazione legata alla scansione ritmica e all'organizzazione tra gli elementi di tipo "diagrammatico". Si tratta dunque di una connessione profonda, ma al contempo fruibile con una grande leggerezza.

³⁸ C. Vernallis, *Experiencing Music Video: Aesthetics and Cultural Context*, Columbia University Press, New York 2004, p. 3.

³⁹ M.B. Korsgaard, *Music Video after MTV: Audiovisual Studies, New Media, and Popular Music*, Routledge, London 2017, p. 73 (traduzione mia).

⁴⁰ Per approfondire il concetto di «figuralità» in semiotica, si veda N. Dusi, *Semisimbolismo, figuralità, tensione in stralci di cinema, tra pittura e musica*, in «Ocula», 9, 2008, pp. 1-10.

Il caso *Afterlife*: varianti narrative “libere”, tra canzone e video

In conclusione, trattiamo il caso del videoclip *Afterlife*⁴¹ (2013) degli Arcade Fire, che si presta come caso studio in quanto la canzone vede associati differenti realizzazioni audiovisive. Per questo diventa un’occasione per riflettere su quali relazioni semantiche vincolino un video a una canzone, ovvero su quanto una canzone sia necessariamente vincolata o meno a uno specifico video.

Il brano è parte dell’album *Reflektor* (2013) per il quale la band ha optato per una campagna promozionale in stile *guerrilla marketing*⁴², con un logo misterioso depositato sui muri di varie città nel mondo, da Londra ad Austin, da Berlino a Roma, e divenuto virale sui social media. In linea con questo approccio comunicativo, anche l’apparato video ha visto una proliferazione di iniziative.

l’album è basato su una sorta di *concept*, o tema conduttore, quale è il mito di Orfeo ed Euridice, a partire dalla copertina, che raffigura l’omonima scultura di August Rodin, ma il tema si ritrova anche nei titoli di alcuni brani. Il mito in questione è connesso a vari sotto-temi, tra i quali il potere dell’amore e dell’arte, e la lotta tra la vita e la morte. La prima apparizione audiovisiva è una sorta di video-album, pubblicato in streaming gratuito su YouTube, associato alle immagini di *Orfeu Negro* di Marcel Camus (1959), un film brasiliano nato da un adattamento teatrale di Vinícius de Moraes e ambientato nelle favelas. Una soluzione inconsueta, prevalentemente promozionale, condotta attraverso una sorta di “libera associazione” tra musica e immagine. Le parole delle canzoni, infatti, scorrono in sovraimpressione in modalità *lyric video*, ovvero come in una sorta di karaoke, senza apparente correlazione con le immagini, come in una sorta di flusso di accompagnamento.

Il passaggio successivo è stato invece la realizzazione del videoclip ufficiale, del tutto scollegato dal film di Camus, con la regia della canadese Emily Kai Bock. Il titolo e le parole del brano *Afterlife* (aldilà) alludono esplicitamente a temi collegati alla morte, ovvero ad un ideale (e onirico) “altro mondo” dopo la vita, così come al lutto e alla sua rielaborazione.

Ascoltando la canzone come semplice traccia audio, si possono desumere vari possibili percorsi interpretativi, ma guardando il videoclip e assieme ascoltando la canzone, le parole e le immagini entrano, per così dire “in risonanza”. Le azioni visibili nel videoclip appaiono manifestare un qualche collegamento alle lyrics, sebbene non in modo esplicito. Se ascoltiamo le parole, possiamo cogliere alcune precise *soggettività* attraverso forme pronominali: un “io”, un “tu” e un “noi”⁴³. Guardando il video, queste forme “attanziali” non sono esplicitamente identificabili in nessuno dei personaggi

⁴¹ <https://youtu.be/EcKinnMXuKg>

⁴² <https://badgerherald.com/artsetc/2013/09/19/arcade-fire-promotion-column-es/>

⁴³ Nella terminologia semiotica, questi processi si definiscono altrimenti come *débrayage* o *embrayage* attanziali, (cfr. A. J. Greimas e J. Courtés, *Dizionario ragionato della teoria del linguaggio*, Bruno Mondadori, Milano 2007) intendendo con ciò la costituzione nel testo di “soggettività” specifiche, attraverso forme pronominali del linguaggio (io, tu, noi, ecc.). Queste forme si ritrovano, in effetti, enunciate nella canzone in questione, ma nel video la possibilità di attribuirle ai personaggi non emerge in modo esplicito.

descritti. Piuttosto, è invece l' "io", che emerge dalle parole della canzone, che si "distribuisce", ovvero si incarna, in ognuno dei personaggi visibili.

Nel video, ambientato a Los Angeles, si racconta di una famiglia messicana, dove il padre, con un lavoro precario, è occupato, dopo la morte della moglie, a prendersi cura dei due figli. Prima dell'inizio della canzone viene inserito un prologo "cinematografico", con dialoghi e suoni diegetici, che si svolge di sera, all'ora di cena. Al calare della notte, poi, inizia la canzone, con il figlio adolescente che si prepara ad uscire, mentre il padre addormenta il bambino. La coincidenza tra canzone e narrazione notturna si spiega con una sorta di *comunicazione onirica* nella quale i tre soggetti sono coinvolti. Ognuno di loro, in sogno, entra in un proprio universo, dove gli accade di incontrare una donna: per il padre la moglie, per i figli la madre.

Nelle lyrics è contenuto uno sviluppo narrativo, articolato in forma poetica, come una sorta di dichiarazione, nella prospettiva soggettiva della *prima persona*. È ciò che può apparire come una sorta di dialogo tra una persona vivente e una persona morta, oppure tra un "io" e un altro assente. Alcuni dei versi qui riportati possono meglio farlo comprendere.

L'inizio della canzone recita:

Afterlife, oh my God, what an awful word

Dunque l'*aldilà* definisce una sorta di campo semantico della canzone, mentre al contempo l'esclamazione "oh my God" fa implicitamente entrare l'*io* nel discorso.

Nel video questa voce rimane come una sorta di *voice over*, una voce narrante non attribuibile a nessuno. Ma nel momento in cui inizia il canto, nel video si vede il *padre*, in un cammino onirico (in bianco e nero), nell'atto di compiere un percorso.

L'*io* narrante della canzone di fronte all'invocazione "Can we", convoca un noi inclusivo:

I've gotta know

Can we work it out?

Let's scream and shout 'till we work it out

Ed è a questo punto che nel video si passa al percorso onirico del *bambino*, che in realtà è un terribile incubo (stavolta a colori), in una lavanderia a gettone, dove il piccolo si vede dapprima trasportato - da una sconosciuta - su un carrello di biancheria sporca, e poi viene chiuso dentro il cestello di un'asciugatrice. Alle parole della canzone: "*Can we work it out? Let's scream and shout 'till we work it out*", il bambino appare muto, disorientato e spaesato.

L'*io* narrante della canzone diventa esplicito mentre dialoga con un *tu*:

Afterlife, I think I saw what happens next

It was just a glimpse of you, like looking through a window

Or a shallow sea

Could you see me?

Qui il video ricalca figurativamente le parole della canzone: ci mostra il padre infilarsi in una casa, e, come recitano le lyrics, guardare da una finestra, vedendo la moglie affacciata da una terrazza sul mare.

La conclusione della vicenda condurrà tutti e tre i personaggi ad una ricongiunzione onirica con il loro oggetto d'amore (moglie o madre), ma ciò che andiamo ad approfondire sono le ricadute sulla relazione tra musica e immagine.

Il video rappresenta una sorta di *possibile pista interpretativa*, costruita a partire dalle parole della canzone, e forse dai temi generali dell'album. D'altra parte, è altrettanto evidente che il video costituisca una costruzione ex-post, e che dunque sviluppi solo uno dei possibili percorsi, tra gli altri.

A riprova di questa ipotesi, viene in aiuto un ulteriore testo audiovisivo, collegato al precedente. Nello stesso periodo dell'uscita del videoclip, a inizio novembre 2013, vi era stata un' apparizione degli Arcade Fire all'interno della prima edizione degli YouTube Music Awards (YTMA). Lo show rappresentava per la band un altro tassello della strategia promozionale dell'album, ma costituisce allo stesso tempo un esperimento innovativo degno di nota in sé.

La ripresa dello show, trasmessa su YouTube⁴⁴ con la regia di Spike Jonze, inizia con una panoramica su un ampio studio al chiuso, con una folta presenza di pubblico, e uno spazio articolato in varie zone. Il presentatore, l'attore comico Jason Schwartzman, dà il via allo spettacolo indicando una zona specifica dello studio televisivo. A questo punto inizia un videoclip con una donna (l'attrice e regista Greta Gerwig) e un uomo nell'atto di lasciarsi, tra sguardi angosciati, lacrime e baci d'addio. L'ambientazione sullo sfondo è quella di un *appartamento*: durante questa sequenza (girata con luci soffuse) non si ha ancora la percezione di una ripresa "live". Nel contempo, mentre inizia il brano *Afterlife*, Greta Gerwig resta sola sulla scena e balla seguendo l'andamento della musica, con scatti e movimenti del corpo che in breve si trasformano in una sorta di possessione ritmica. Presa da una danza energetica, costei esce dall'appartamento mentre sullo sfondo risuonano le parole della canzone: "*Can we work it out? Let's scream and shout 'till we work it out*". La chiave interpretativa, fin qui desunta a partire dal videoclip ufficiale di Emily Kai Bock, è totalmente mutata.

Sempre a passo di danza, si entra in uno *spazio teatrale* composto da un bosco notturno dall'aspetto artificiale, attraversato da neve cadente. In realtà, come si può vedere guardando con attenzione lo show, l'intero videoclip è girato nello studio, il quale è stato diviso in zone separate, con differenti scenografie.

L'attrice si sposta da una zona all'altra, all'interno del medesimo studio, mentre nello spazio antistante, la band degli Arcade Fire sta eseguendo il brano dal vivo. Dunque ciò che stiamo vedendo è solo apparentemente girato con lo stile di un videoclip, ma in realtà si tratta di una performance live.

Si odono dopo poco, infatti, le prime urla del pubblico presente in studio, il quale sta osservando da vicino la scena nell'atto di essere filmata. Da questo momento in poi, lo *spazio diegetico* della finzione narrativa inizia ad incrinarsi. Improvvisamente, alle spalle di Greta appare il cantante della band (infilatosi nella scenografia) mentre, con in mano il microfono, esegue il brano dal vivo. La finzione narrativa è stata a questo punto

⁴⁴ <https://youtu.be/yXBPbnv1H-U>

contaminata da uno spazio video-performativo che, se da un lato non ha più nulla di finzionale, dall'altro esibisce un improvviso "sguardo in macchina", operato da Greta in modo reiterato. Non siamo più nella finzione cinematografica, bensì siamo nel *contratto comunicativo del videoclip*, costruito per "guardare negli occhi" il fruitore. Ma qui avviene un'ulteriore rottura delle convenzioni: Greta scende dal bosco teatrale e corre lungo una passerella a livello del parterre, seguita da un gruppetto di giovanissime ballerine che, attorniate dal pubblico, saltano e ballano assieme a lei. Nel finale il video è diventato un *concerto*: una sorta di bagno di folla in cui band, ballerine, attrice e pubblico danzano assieme al ritmo della musica.

In queste brevi sequenze il regista Spike Jonze ha messo assieme, e fatto collassare tra loro, molte delle convenzioni e delle soglie che caratterizzano il linguaggio del videoclip. Ripercorrendo i vari passaggi, vediamo:

- dapprima una *convenzione narrativa* tradizionale, con una finzione dotata di "quarta parete", salvo il fatto che i personaggi non parlano, ma sono in qualche misura guidati dallo sviluppo della canzone;
- poi, quando la protagonista decide di agire, avviene una trasformazione passionale, e così dal dramma si passa al *musical*: l'attrice si muove e balla al ritmo della canzone, che a questo punto diventa il filo conduttore della narrazione.
- successivamente si passa a un momento di *transizione* nello spazio "teatrale" del bosco, che si rivela essere uno stato performativo ancora embrionale, nel quale appare il leader della band.
- qui Greta Gerwig può entrare in scena e ricevere la sanzione positiva del *pubblico*, riunendo assieme la diegesi *finzionale*, quella *musicale*, e quella *performativa*.

Dunque non solo sono state smascherate e messe in luce molte delle convenzioni del videoclip, ma questo caso ci dimostra come la relazione tra immagine e musica nel videoclip sia molto "labile", e per nulla basata su una reciproca dipendenza. Una canzone può essere associata a diverse tipologie di videoclip, costruendo narrazioni tanto diverse, quanto potenzialmente del tutto slegate tra loro.

Come a voler dire che da una canzone si possono costruire infinite varianti di videoclip, e che i contratti in queste tipologie testuali, potremmo ironicamente dire, siano tipici di "coppie aperte", ovvero di canzoni che possono avere rapporti con più video contemporaneamente: un lyric video, un videoclip tradizionale, e uno show live, possono valere in tutti i casi come testi in grado di sviluppare alcune delle potenzialità di una canzone.

Conclusioni

Analizzando il linguaggio del videoclip nel suo complesso, e in specifico nella relazione tra suono e immagine, emergono alcune specificità testuali di questo formato mediale, le quali non sono state ancora sufficientemente approfondite. Il videoclip, nato in modo quasi spontaneo, come fenomeno di tendenza e senza una particolare dignità artistica,

in realtà si dimostra uno strumento espressivo efficace, che perdura nella contemporaneità. La sua natura composita e sincretica evidenzia come il linguaggio audiovisivo presenti convenzioni e prassi complesse, e in parte “nascoste”, le quali, oltre a fornirci squarci di comprensione rispetto al fenomeno specifico del videoclip, si connettono a questioni teoriche di ben più ampia portata, dal punto di vista estetico e semiotico, legate al cinema e all’audiovisivo in genere.

Per la primarietà assegnata al suono rispetto all’immagine, il videoclip mostra una certa unicità nel panorama audiovisivo, e al contempo pone interrogativi sui processi che legano il suono all’immagine. Inoltre, dimostra una propria specifica capacità narrativa che infatti ha influenzato, nel passato come nel presente, le modalità del racconto mediale *tout court*. Nonostante ciò, la letteratura scientifica dedicata al fenomeno si dimostra carente, e pone un problema più generale sulla necessità di costruire una più articolata *teoria generale dei processi di linguaggio audiovisivi*, che vada al di là delle singole specificità espressive (cinema, televisione, media digitali, videoarte, ecc.).