calculable manufacturings of pleasurable musical material by way of audio file manipulation. Distinctive concepts and formats of musicalisation are examined in view of their underlying compositional ideas and the affective stimuli they pass on. As musicalisations of found objects on YouTube contributes to the normalisation and perpetuation of pleasurable musical engagement with pre-recorded realia, issues of "indexploitation," surveillance, and ridicule shall be problematised in this context.

Computational surface and interface effects represent a third category of media objects that inform vernacular re-composition in the YouTube era. As signs, signals, and surfaces effected by computational interfaces represent banalities of networked interaction, their re-imagination, representation, and re-production in co-creative musical practices affords communicative connectability and invokes a certain vernacular ethos. The analysed implicitly or explicitly interface-centric practices are characterised by aesthetic approaches concerned with audiovisual patterns, concrete representations and re-compositions of graphical user interfaces, or computationally attained flaws and failures. As a result, computational surface and interface effects are re-domesticated within practices of musical produsage on the platform, unfolding their very own meta-referential potentialities.

In general, within largely disembodied and increasingly globalised online communities, communal cohesion and self-narration evolve largely based on the implicit and explicit meta-reflection on the status of "being online," the remediated digital and pre-digital cultural objects, and the (simulacric) cultural memory and compositional repertoire shaped through ongoing produsage. Thus, the aspect of meta-referentiality, which is indispensable for my overall argumentation, will accompany and, at times, guide the following categorisations of media objects of fascination and the practices of re-composition evolving around them.

3.1 Meta-Discursive Music Video

The shift from music video's domestication on music television – following the launch of MTV in 1981 – towards its digital production and distribution has caused an ongoing discussion about what today, in digital environments characterised by accelerated and ubiquitous processes of media convergence, actually *defines* a music video. The reason for the confusion lies in the ramifications of music video's emancipation from its once seemingly "natural" media environment of music television. Obviously, the fundamental quality of music video remains intact in the age of digital distribution: "[a]ny music video operates by visually remediating music (recasting a preexisting song visually), but also by musically remediating the image (structuring the image according to musical logic)."[1] Moreover, one could say that the remediating effect of music videos on *other* media has not changed, as they still adopt and juxtapose various sources, as Carol Vernallis points out with regard to appropriations of – or allusions to – aesthetic objects from various media such as film and commercials.[2] Mathias Bonde Korsgaard argues that, historically, "the music video aesthetic was itself formed on the basis of remediating other medial expressions, and, in turn, the aesthetics of music video have arguably been fused back into these medial expressions from which it originated."[3] On a large historical scale, these processes of remediation between music video and other (audiovisual) media can usually be imagined as reciprocal; however, while, for instance, the mutual impact of audio-visual remediation in music video and cinema could historically be traced and exemplified, today's media convergence brings about an entanglement of media that makes it difficult to draw the lines between individual media, since media-specific delivery channels and technologies have become largely obsolete. This is also the case with regard to new forms of interactive music videos, which have emerged in awareness of what Paul Hearsum called a "24-hour *multiple* functionality (to listen, to watch, to select, to read, to add comments, and to participate)" compared to MTV's "*dual* functionality (to listen and to watch)."[4] In the interactive video for Bob Dylan's song "Like a Rolling Stone," which promoted the release of Bob Dylan's complete CD Box set in 2013, the user can flip through 16 differently themed "TV channels," each representing a program (shopping, cooking, news, sports, reality TV, etc.) that showcases a different character lip-syncing Dylan's song.[5] The video is only one example for the hybrid manner in which media evolve within a convergent media landscape. Other artists spawn fan contributions, for example by deliberately integrating animating choreographies in their music video performances, thereby encouraging the widespread creation and circulation of fan imitations in social media. Psy's rider dance in his music video for "Gangnam Style" (2012), which prompted a stream of imitations all over the world, is arguably one of the most well-known examples.

---

[1] Andrew Goodwin, *Dancing in the Distraction Factory: Music, Television and Popular Culture* (Minneapolis: University of Minnesota Press, 1992), 509.

[2] Carol Vernallis, "Music Video's Second Aesthetic?," in *The Oxford Handbook of New Audiovisual Aesthetics*, eds. John Richardson, Claudia Gorbman, and Carol Vernallis (Oxford: Oxford University Press, 2013), 459.

[3] Mathias Bonde Korsgaard, *Music Video after MTV: Audiovisual Studies, New Media, and Popular Music* (London and New York: Routledge, 2019), 144.

[4] Paula Hearsum and Ian Inglis, "The Emancipation of Music Video: YouTube and the Cultural Politics of Supply and Demand," in *The Oxford Handbook for New Audiovisual Aesthetics*, eds. John Richardson, Claudia Gorbman, and Carol Vernallis (Oxford: Oxford University Press, 2013), 491.

Generally, such tendencies towards interactivity and participation can be seen as both appropriations of functions associated with new digital media as well as reconfigurations of the forms, formats, and functions of music video. Entire "music video albums" as well as gamified music videos with non-linear narratives bear witness to a new-found temporal malleability and discontinuity of "post-televisual" music videos,[6] as Vernallis points out: "[…] music videos no longer have to fit the short lengths of pop songs, or present them without interruption […]: we may be able to say only that a music video is a relation of sound and image we recognize as such."[7] Her very broad definition of music video is not only informed by new corporate adaptions of principles of digital and social media, but also – and, in the context of this chapter, more importantly – by the dissolution of the dichotomy between production and consumption in the context of vernacular audiovisual (re-)composition in social media. The boundaries of music video have become blurred in the digital sphere, as the medium is decoupled from its former delivery channels and, in a media ecology characterised by convergence and participation, cannot be exclusively defined by its conventional function as an auteurial promotional vehicle (for an artist, a song, an album, or for a film which features the artist's music) anymore. Entangled within an environment of networked consumption, vernacular music video composition follows a participatory logic and becomes involved in processes of media hybridisation. Any audiovisual artefact, following its upload to a video platform like YouTube, becomes easily attainable modular material for user-generated music video-like creations – or for remixes, mashups, and re-enactments of pre-existing music videos. Remediated by the digital principles of modularity and variability, music video is thus embedded in an environment of continual re-appropriation and re-iteration and, consequentially, expands its own textual affordances. Of course, music video has always potentially served as a "producerly text," insofar as it, due to its inconsistencies, its references to other media texts, and its often enigmatic character, encourages creative readings and productive engagement with the offered audiovisual texts – an example for this would be concomitant meta-narratives of stardom that develop or shift according to the interrelations of artistic production, medial (self-)representation, and fan discourse. However, enabled by the quantity and accessibility of audiovisual material on video platforms as well as the affordability and suitability of video editing software for beginners, producerly behaviour towards music video now visibly manifests and circulates in new forms and formats of audio-visual remediation, drawing large audiences and encouraging further contributions. Regardless of whether we potentially define any form of reciprocal remediation of music and images as a music video or resort to describing new emergent hybrid configurations as "music video-like" or "post-music video,"[8] a drastic quality shift can be noted which is brought about by the entry of music video as a media artefact into the domain of Internet-mediated vernacular content creation: while any conventional music video is fundamentally embedded in cross-media reference and processes of remediation – and thus entangled in a transtextual and transmedial fabric – and exclusively "devotes itself to referring to other images, other narratives" without necessarily taking account of its implicit self-referentiality,[9] vernacular re-appropriations of music video-related audiovisual figurations are characterised by the users' awareness of their derivativity as well as of the surrounding framework of networked consumption they are situated within and thus potentially elicit a high medium-awareness in the recipient. Especially mashup techniques can be conceived of as a performative practice which, by way of hybridisation, can generate an awareness of the involved technologically and culturally significant medial systems. According to art and media theorist Pamela Scorzin,

---

[5] Bob Dylan, "Like a Rolling Stone," November 2013, interactive music video, https://eko.com/v/like-a-rolling-stone?autoplay=true.

[6] For this term, see Korsgaard, *Music Video after MTV*, 17.

[7] Carol Vernallis, *Unruly Media: YouTube, Music Video, and the New Cinema* (Oxford: Oxford University Press, 2013), 27.

> "mashup" in its modern sense of a recent media phenomenon, a resampling of previously existing medial or artistic material, can currently be consid-

---

[8] See Korsgaard, *Music Video after MTV*, 173–196.

[9] Angela McRobbie, *Postmodernism and Popular Culture* (London: Routledge, 1994), 17.

> ered the most popular source of metaization in contemporary culture; consider, for instance, a digital media file containing text and/ or graphic, audio or video elements as well as animations, thus recombining, recoding, restructuring and modifying pre-existing digital material in order to create a derivative and hybrid new work.[10]

"Metaisation" is a term coined by Werner Wolf to denote transmedial processes of "usually non-accidental self-reference produced by signs or sign configurations which are (felt to be) located on a logically higher level, a 'meta-level,' within an artefact or performance." The resulting meta-reference is described by Wolf as a form of "self-reference, which can extend from this artefact to the entire system of the media, forms or implies a statement about an object-level, namely on (aspects of) the medium/system referred to. Where metareference is properly understood, an at least minimal corresponding 'meta-awareness' is elicited in the recipient."[11] As Henry Jenkins states with regard to the evolution of media in an environment of media convergence, "once a medium establishes itself as satisfying some core human demand, it continues to function within the larger system of communication options."[12] YouTube's communication options, afforded by the platform's "multiple functionality," encourage a discursive engagement with the mediality, textuality, and materiality of music video itself. Mashups, parodic imitations and re-enactments of concrete or generic audiovisual material thus represent polytextual methods that subject music video – as well as the creator's stance – to communal discourse by way of meta-reference and self-reference. The following analyses focus on aspects of interplay between bodily performance, visuals, language and narrative, musical references, and sonic modifications in order to shed light on platform-specific forms, formats, and functions of vernacular music videos "in the second degree" and the discursive levels of their inherent metaisation.

Fan Videos as Communal (Meta-)Discourse and Self-Reference

Of course, the history of music video-like amateur creations does not begin with the emergence of online communities. In fact, it goes all the way back to

---

[10] Pamela C. Scorzin, "Metascenography: On the Metareferential Turn in Scenography," in *The Metareferential Turn in Contemporary Arts and Media*, ed. Werner Wolf (Amsterdam: Rodopi, 2011), 261.

[11] Wolf, "Metareference across Media," 31.

[12] Jenkins, *Convergence Culture*, 14.

media fandom of the 1970s. In a practice known as "vidding," fans engage with the combination of video footage and music from mass media sources, usually television broadcasts or film productions. As numerous studies underline, vidding community has, from its beginnings, been female-dominated. The role of gender in vidding has been extensively theorised in the field of fan studies. Usually, the practice of vidding is conceived of as a response to male-led professional media production, driven by the co-creation and passing on of desire events on the basis of communally established generic modifications to pre-existing narratives.[13] Whereas early vidders created montages by recording, selecting, cutting, and recombining televised material using VHS machines, nowadays video platforms and editing softwares provide the archival and technical means for vidding. Vidding is a narrative practice aimed at highlighting, developing, or subverting one or multiple source texts – usually from series or movies – by setting selected audiovisual segments to music. The outcome ranges from reverential homages that merely summarise certain moments or aspects of a storyline to deliberate re-contextualisations of images that tell alternative storylines or focus on the modification of relationships or character developments.

Although vidders often stress the differences of their practice to conventional music videos, their historical, medial, and aesthetic backgrounds are related, as Angelina Karpovich notes:

> Having emerged at roughly the same time, and sharing a similarly complex relationship with the media of film and television, music videos and fan videos have more in common than has previously been acknowledged […]. Fan videos are […] the result of their creators' cumulative media experience, in which aesthetic and conceptual influences from a variety of texts, genres and forms result in multi-layered texts that are steeped not only in the immediate context of the source media text, but in the very same audiovisual aesthetic that has also produced, among other forms, the music video.[14]

---

[13] See, for example, Katharina M. Freund, "'Veni, Vidi, Vids!' Audiences, Gender and Community in Fan Vidding," (PhD thesis, University of Wollongong, 2011), 2–3, https://doi.org/10.7916/d8-6rte-j311. Moreover, in her historical account of the development of vidding from the 1970s to the present, Francesca Coppa accentuates the queer and critical dimensions of vidding. See Francesca Coppa, *Vidding: A History* (Ann Arbor: University of Michigan Press, 2022).

[14] Angelina Karpovich, "Reframing Fan Videos," in *Music, Sound and Multimedia: From the Live to the Virtual*, ed. Jamie Sexton (Edinburgh: Edinburgh University Press, 2007), 27.

Like in music videos, the images of fan videos – or simply: vids – are cut to a pre-selected song, resulting in deliberate synchronisms between music and video as well as in moments of non-synchronicity. However, whereas Michel Chion, with respect to music videos, observes a potential liberation of the image "from the linearity normally imposed by sound" and points to the usually broken or non-existing narrativity of music video's images, moments of non-synchronised "visual polyphony" in fan videos must be evaluated differently due to their essential narrative function. Unlike the imagery of music videos, in fan videos, the visual layer is fundamentally anchored by the chosen song and never decoupled from, or indifferent to, the overall audiologovisual textual work. For example, in the fan video "I hate you, I love you // Johnlock," @Johnlocked explores the relationship between the main characters Sherlock Holmes and Dr. John Watson from the British television series "Sherlock" by setting video segments from the series to the song "I hate you, I love you" by Gnash ft. Olivia O'Brien.[15] Synched to the opening piano chords and rhythmically accelerating over the course of the video, a prism of sequences showing John and Sherlock in confrontational, thoughtful, and affectionate moments are used to sketch a fictional relationship drama between the two. Throughout the video, Olivia O'Brien's voice is used to portray Sherlock's inner feelings towards John; Gnash, in return, sings "on behalf of" John. The absence of both singers in the fan video helps avoid the double address characteristic for music video, in which the "personality of the storyteller usually overwhelms characterization within the story," entailing a "conflation of the real/implied authorial voice."[16] According to Henry Jenkins, in fan videos "[t]he performer's personality must be effaced so that the singer may speak more effectively on behalf of the fictional character."[17] While the video is indifferent towards the singers' personalities, the music and the lyrics are highly empathetic in regard to the images. Beyond the musical expression of "its participation in the feeling of the scene," a fundamental quality for what Michel Chion calls "empathetic music,"[18] the singers add emotional value as quasi-narrator's voices, aiming at a semantic listening of the song. The lyrics are used to anchor the images and create a shift in the paradigmatic dimension of the narrative, in that the audiologovisual text of the video "takes character and settings and makes a non-temporal sense of them" beyond the syntagmatic narrative dimension of temporality, cause, and effect.[19] The concrete interplay of music, lyrics, and selected sequences in @Johnlocked's video represents Sherlock's and John's relation as a homoerotic relationship full of longing and heartbreak, irrevocably obstructed by John's marriage with Mary Watson. In the first chorus, for instance, Sherlock (through the voice of Olivia O'Brien), accompanied by images of them arguing and laughing together, confesses his complicated feelings towards John ("I hate you, I love you, I hate that I want you"). At the words "you want her, you need her and I'll never be her," scenes of Mary and John hugging and dancing in their wedding dress and suit are followed by a close-up shot of Sherlock's sad face, seemingly looking at them. After an answer by John (Gnash), telling him that "love and trust are gone" – probably due to Sherlock's egocentric behaviour throughout the series, the next chorus is sung as a duet and ends on a held piano chord over which Sherlock's original speech from the series resonates: "It's always you, John Watson, you keep me right." The video ends with the conclusive line from the chorus ("I'll never be her") and a scene of Sherlock leaving the wedding party of John and Mary.

---

[15] @Johnlocked, "I hate you, I love you // Johnlock," November 26, 2016, YouTube video, 3:46, https://www.youtube.com/watch?v=Tvx5s1dhbDE.

[16] Goodwin, *Distraction Factory*, 76.

[17] Jenkins, *Textual Poachers*, 235.

[18] Chion, *Audio-Vision*, 8.

The anchorage of the images through the lyrics guides the viewer through the signifieds of the video: The scene of Sherlock leaving the party, for instance, signifies his inner parting from the idea of a romantic relationship with John. However, contrary to conventional music videos, every sign in the video, as "the associative total of a concept and an image,"[20] is secondary to the mythical system it is caught within. According to Roland Barthes, "myth is constructed from a semiological chain which existed before it: it is a second-order semiological system."[21] On the one hand, the fan video by @Johnlocked builds on pre-existing fan theories about Sherlock's (and John's) homosexuality, fuelled by Sherlock's seeming lack of sexual desire for women and assumptions by other characters of the series – for example by their landlord Mrs. Hudson, who mistakes them for a couple (a regularly returning running gag throughout the series). On the other hand, the genre of vidding itself is characterised by generic modifications to the paradigmatic dimensions of narrative. For instance, the "slash" subgenre, which the video belongs to, focusses on creating connotations of a homosexual bond between two usually heterosexual characters from a fictional media universe. Against this backdrop, the audiologovisual textual work of @Johnlocked's video becomes mythical, in that it forms a "*metalanguage*, […] a second language, *in which* one speaks about the first."[22] In mythical approaches to narratives, every sign within the linguistic system becomes a signifier of a "greater system which it builds and of which it is only a part."[23] These second-order signifiers that lend themselves to myth do not create meaning, as "the meaning is *already* complete,"[24] but function as *form*, as mythical signifiers emphasising the "cultural-ideological system that underlies the syntagmatic flow of the narrative," as John Fiske puts it.[25]

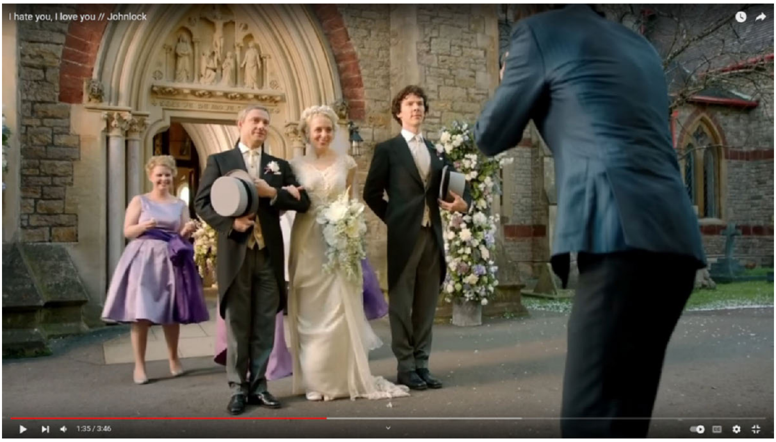
---

[19] Fiske, *Television Culture*, 130.

[20] Barthes, *Mythologies*, trans. Annette Lavers (New York: Noonday Press, 1972), 113.

[21] Ibid.

Figure 1: Sherlock as the fifth wheel, standing next to the newly-wed couple.



Still from @Johnlocked's video "I hate you, I love you // Johnlock" (2016).

The produced text in vidding practices is built on the symbolic potentiality of pre-existing texts from media productions and franchises and evolves transmedially as an open, off-centred structure based on participation and generic symbolic play. With his remark that text "only exists in the movement of a discourse" and necessarily "knows itself as text,"[26] Barthes hints at the fundamental implicit meta-referentiality of textual reception-as-production. Today, the conflation of writing/production and reading/consumption *materially* manifests in manifold ways as a shared repertoire of fannish (meta-)texts and (sub-)genres, perpetuating the communal generation of producerly text, which "relies on discursive competencies that the viewer already possesses."[27] At times, the meta-discourse of fan videos even contains direct self-references to the vidding community: In the video "Us," @lim creates a homage to media fandom itself by setting images from 34 different series and movies to the song "Us" by Regina Spektor.[28] Containing lines like "we're living in a den of thieves, rummaging for answers in the pages […] and it's contagious," the lyrics represent the self-understanding of a creative community focussed on practices characterised by the re-contextualisation and de-centralisation of media texts, while hinting at the unavoidable copyright issues linked to remix culture.[29]

---

[22] Barthes, *Mythologies*, 114.

[23] Ibid., 113.

[24] Ibid., 116.

[25] Fiske, *Television Culture*, 136.

DIY Ethos and Aesthetics of Profanity

Beyond the remediation of music video aesthetics as a vehicle for meta-referential transmedial storytelling and communal self-reference, vernacular formats of music video-like composition are focussed on the medial texts and surrounding discourses of (commercial) music video production. Re-enactments of music video come to mind first, as they represent a long-standing approach to metaised music video from pre-digital times. For instance, Weird Al Yankovic's parodies of iconic songs and music videos such as "Eat It" (Michael Jackson – "Beat It"), "Smells Like Nirvana" (Nirvana – "Smells Like Teen Spirit"), or "Amish Paradise" (Coolio – "Gangsta's Paradise") have been broadcast by music television channels like MTV or VH1, propelling sales of his corresponding parody albums and earning him five Grammy awards between 1985 and 2019, one for "Best Concept Music Video" for his 1989 parody video of Michael Jacksons "Bad" ("Fat"). Compared to this rather conventional embedding of music video parody within value chains, as a promotional vehicle for the artist and their output, YouTube affords the upload and circulation of do-it-yourself music video parodies largely unaffected from pre-digital industrial gatekeeping and economic constraints. Jon Lajoie's video "Show Me Your Genitals" is a good example for a parodic music video carried by a no-budget production and a corresponding DIY ethos. Lajoie started filming and uploading his own music videos in 2007, featuring himself embodying various characters. The video for "Show Me Your Genitals" only shows a graffitied wall, a car, and Lajoie in his role as "MC Vagina," wearing a baseball cap and a colourful vintage shirt tucked into his shorts. The video editing is characterised by automated effects like rudimentary and demonstratively amateurish text animations which highlight the song's lyrics. The backing track is reduced to a four-bar drum machine loop, consisting of elements one could find in a proper hip-hop beat: a syncopated bass drum, 16th notes on the hi-hat, a clap sound for the snare, and the use of triangle samples. The poor quality of the samples and the loop's redundancy, in combination with Lajoie's outward appearance, are the first perceptible constituents of the overall performance of failure that characterises the whole video. The de-familiarised visual and musical associations to song and video productions in the hip-hop genre provide the backdrop for the lyrics, which, in their blatant sexism, aim at rendering automatised and implicit misogynistic (and homophobic) tropes in rap lyrics explicit:

---

[26] Barthes, Barthes, "From Work to Text," 157.

[27] Fiske, *Television Culture*, 95.

[28] @lim, "Us | Multifandom," June 2, 2007, YouTube video, 3:55, https://www.youtube.com/watch?v=_yxHKgQyGx0. As the title suggests, videos with references to several media sources are called "multifandom."

[29] Moreover, the video symbolically refers to academic discourse about media fandom by including footage of a talk by Henry Jenkins at a conference of the "Converge Culture Consortium."

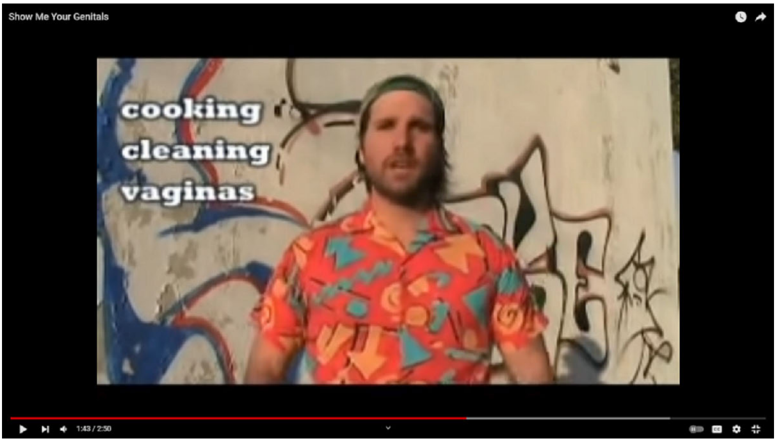
> […] I wanna see your bum, I don't care what you say,
> No, I don't have feelings 'cause feelings are gay,
> Something something in the month of May,
> Bitches love my penis 'cause it's really big […][30]

Throughout the song, MC Vagina stumbles through verses that include awkward rhymes or ignore the rhyme scheme altogether. His pointedly unskilful rap performance is accompanied by the scenery of the video: a graffitied wall and an old car without a crew or decorating women that would bolster up his musical or bodily performance. Clichéd symbols of masculinity, objectifying depictions of women from a male gaze, and conventions of self-staging in hip-hop videos become noticeable exactly through their absence or deformation. The result is a lack of credibility with regard to the lyrics, which is underlined by the song's awkward outro: "Ahh yeah, that's right, shake your… bums. I'm out of here… I gotta have sex… with a lot of girls."[31] The appearance, utterings, and demeanour of Lajoie's persona MC Vagina are staged as an "as-if-authentic" performance, serving as a vehicle for a parodic commentary that lets meta-narratives and phantasies of stardom collapse and, by way of exaggeration, points to interchangeable, and thus often uncontested, misogynistic tropes regarding the visual and semantic texts of hip-hop productions.

---

[30] @Jon Lajoie, "Show Me Your Genitals," May 31, 2008, YouTube video, 2:50, https://www.youtube.com/watch?v=qqXi8WmQ_WM.

Figure 2: Still from Jon Lajoie's music video "Show Me Your Genitals" (2008).



When MC Vagina raps "it's not sexist 'cause I'm saying it in a song," the lyrics even point directly to the issue of automatisation and banalisation of sexism in an ironic manner.[32] Overall, Lajoie's video fits the characterisation of parody by Linda Hutcheon, who describes parody as a "bitextual synthesis," stating that, "[w]hile the act and form of parody are those of incorporation, its function is one of separation and contrast."[33] Conventional forms of audiologovisual figurations in rap videos, for example popular motifs of visual and linguistic anchorage, are contrasted by an overall aesthetic of failure and a semantic strategy of rendering genre-typical tropes noticeable by way of ironic hyper-affirmation. The result is an implicit meta-reference to generic music video production beyond the parodic uptake of a single video or a distinctive personal style, characterised by a stronger bitextual determination and a wider meta-referential scope than could be attained by methods of direct quotation and imitation. Before the emergence of online video hosting, a music video parody which functions as an implicit meta-text on overall conventions of music (video) production without referring to a concrete popular source text would have been deemed incongruous with industrial value chains optimised for pre-digital mass media. More importantly, however, what sets many DIY music videos apart from any industrial standard is the aim of achieving an "unruly" aesthetic. Lajoie's video serves as a good example in its demonstrative profanity, brought about by lo-fi means of production and a DIY ethos, an overall performance of failure, and a lack of recognisable quotations – entailing a potentially perceived lack of palatability as well as issues of readability and discursivity resulting from its consequential hyper-affirmative and ironic character. On YouTube, vernacular music video parodies have become established as a potentially spreadable – and therefore profitable – format, enabling Lajoie to start a popular YouTube channel featuring various no-budget parody videos, garnering over a million subscriptions.

---

[31] Ibid.

[32] Ibid.

[33] Linda Hutcheon, *A Theory of Parody: The Teachings of Twentieth-Century Art Forms*, 2nd ed., (Chicago: University of Illinois Press, 2000), 33–34.

Besides the spread of DIY music video parodies, the biggest change that comes with networked music video (co-)creation is the engagement with and continuous re-interpretation of *pre-existing* music videos with discursive and meta-referential implications. For instance, under the genre label "shred," re-dubs of music videos have emerged in the early years of YouTube and, since then, become the most popular form of music video-like produsage based on techniques of remix. The first video labelled as a "shred" was uploaded in 2008 by Santeri Ojala (or @StSanders),[34] who is also commonly regarded as the inventor of the genre. Shreds are characterised by the synchronisation of added audio to a muted pre-existent video clip of a musical (live) performance. In fact, the synchronisations in shreds are very literal, as usually mostly the visible events in the video are sonically re-interpreted. For example, in Ojala's video "Eddie Van Halen Shreds," a live guitar solo by Van Halen, meticulously

---

[34] @StSanders, "Kiss Shreds," November 7, 2008, YouTube video, 4:09, https://www.youtube.com/watch?v=Kw5oJoUYTb8.

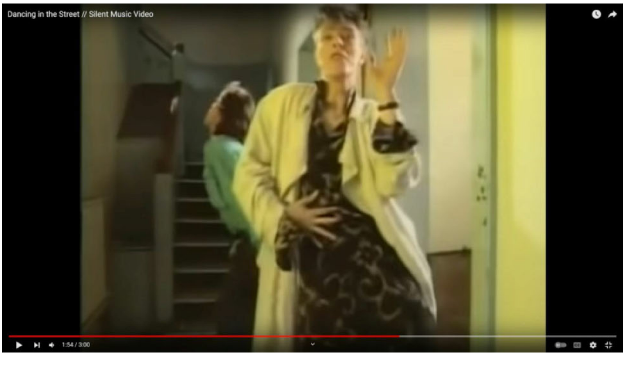
dubbed to seemingly fit his movements, turns into a grotesque display of dilettantism. The musical outcome clashes with the visual layer: Van Halen's repertoire of poses and playing techniques are juxtaposed with a musically nonsensical dub track that bears no resemblance to the original audio. Since the other band members' playing is only dubbed when they are actually on display, Van Halen's solo seems all the more out of place, as the effects of awkward pauses, pickup noises, and cacophonic playing, which are at odds with his bodily performance of musical (and masculine) virtuosity, become enhanced. Besides this approach of sonic reduction, a few sounds are also added beyond the visual layer to ironically comment on the musical performance, such as the added – rather reserved – applause from the audience after multiple segments of the failed solo. The video is a good example for the way in which incongruencies between audio and video are created to elicit a meta-awareness in the recipient, as it can be read as a meta-referential contribution countering the ubiquitous display of virtuosity, especially in guitar-related content on YouTube and beyond. In fact, most early shreds on YouTube focussed on ridiculing performances of famous persons "shredding" their guitars – hence the label "shred." In addition to blatantly nonsensical shreds, there is a widespread ethos of creating "convincingly bad" dubs which are so subtle they might lead the recipient to mistake the video for the original. This is often the case in shreds that closely imitate the original audio track and integrate only subtle rhythmic and melodic flaws.[35] Yet, by way of re-synchronising the audio with the visual layer, which results in a performance of failure, shred videos, no matter the concrete source text or approach to dubbing, always create a decodable audio-visual dissonance that potentially renders normalised and conventional performances and significations of virtuosity, musical expertise, stardom, and gender visible – if the recipient stumbles upon the cracks in the performance and produces their own critical subtext.

Notwithstanding these considerations, a much more concise form of shred has been realised by the transferral of shredding practices to commercial music videos in the format of so-called "musicless music videos" (alternatively called "silent music videos" or "music videos without music"), which are characterised by the absence of the original music and the addition of synchronised audio dubs, using only "as-if-diegetic" sounds. Musicless music videos arguably bring about a more "limited reading," as the object and frame of reference can be easily inferred – and this is not solely afforded by the denotative paratext of the videos' titles: The absence of (non-diegetic) music strips the video of its main function as a tailor-made visual composition to a pre-existing song and, in doing so, consequentially elicits a high meta-awareness of formal aspects and patterns of conventional and automatised audiovision in music video productions. It is safe to say that the re-interpretation and re-contextualisation of the visual layer by way of dubbing techniques bears more meta-referential gravity in musicless music videos compared to conventional shreds, as the music often is entirely missing and the recipient inevitably finds themself confronted with an unbridgeable gap between their aesthetic expectations towards commercial music videos and the concrete audiovisual outcome. The radical synchronisation of on-screen action and diegetic sounds creates an acoustic space full of aural codes that denote a live situation, laying the "visual polyphony" and broken narratives of the video bare and entailing a kind of "de-naturalisation" regarding the (non-)diegetic qualities of music and sound in original commercial music videos.

---

[35] This is the case in @magetimusic's 9-minute (!) shred of Miles Davis' "So What." The video starts as a slightly sloppy rendition of the piece and gradually descends into a more and more uncoordinated musical situation, with bandmates cueing each other (and missing their cues), unrelated musical quotes in the solos – for example the theme from "Spiderman" – and the sonic appearance of strange keyboard patches around the six-minute mark which, in the absence of keyboards in the visual scenery, finally breaks any illusion. See @magetimusic, "Miles and Coltrane Shred," January 6, 2010, YouTube video, 9:23, https://www.youtube.com/watch?v=9ZCz--kHbTs.

Figure 3: Mick Jagger and David Bowie dancing. Still from @Strack Azar's musicless music video of "Dancing in the Street" (2016).



@Strack Azar's shred of David Bowie's and Mick Jagger's video for their duet cover of "Dancing in the Street" is a good example for the de-familiarising effect of musicless music videos.[36] In the video, which was filmed at night, Bowie and Jagger can be seen dancing and singing in a warehouse, a staircase, and a street. Instead of the original audio track, an acoustic space of silence is introduced by the sound of clittering crickets, which, later in the video, are accompanied by owl calls and, as indicators of urban life, distant police sirens and gunshots. Sounds illustrating the bodily movements of the two singers, like stomping feet and the rustling of their clothes, add to the overall sonic diegesis. Musical passages from "Dancing in the Street" are only introduced by lip dubs following the melody and lyrics of the song synchronously to the lip movements of Bowie and Jagger, albeit in a comedic fashion: the soft and shaky singing, unsupported by an instrumental backing track and interspersed with voice cracks, creates a stark contrast to the expressive and exuberant dancing, which seems completely out of place now. By closely following the visual layer, @Strack Azar inverts the original audio-visual relation of the music video, highlighting the sudden discrepancies of quick cuts and unmediated video segments in relation to the now-diegetic audio, as they are stripped of their function as "visually polyphonic" elements due to the lack of an underlying song that would serve as a background for rhythmic and prismatic video editing. For example, a short segment of David Bowie jumping in slow motion causes a surprising interruption, as @Strack Azar's synchronisation strictly follows the image and disrupts the prior lip dub of Bowie's singing by accompanying his jump with a loud scream ("Dancing in Chic-aaaaaaaargh!"). Overall, the diegetisation by use of Foley and lip dubs in musicless music videos entails the dominance of the image over the audio layer, reducing the logic of music video to absurdity. In dissecting the conditions of audio-vision in commercial music video productions by way of re-functionalisation of the interplay between music, sound, and video, the shred subgenre inherently unfolds a transformational impact that goes beyond concrete source material and elicits a critical meta-awareness of the aesthetic object of music video – and its production – as such.

As the examples show, the "vernacular ethos" of YouTube-situated music video-like contributions may, on the one hand, arise from the place the videos are coming from, in that they are self-produced by individuals using easily available technological means (often without any budget). Secondly, their content performatively invokes attributions of non-institutionality. In the case of music video-like fan vids, this results from 1) a discursive orientation towards networked spaces of affinity and 2) the remediation of vernacular creativity through forms and formats based on the narrative re-imagination and re-appropriation of popular media texts by way of generic audiologovisual metalanguage. Beyond the self-referred non-institutional stance of vidders, the music video parody by Jon Lajoie and @Strack Azar's musicless music video paradigmatically highlight the potentials of deconstructing aspects of industrialised music video production, including meta-narratives of stardom, clichéd verbal and bodily performances, or conventions of audio-vision in commercial music video productions. In summary, the vernacular of music video-like media composition is invoked performatively in contrast to institutionalised practices of commercial music video and multimedia production, opening up a field of produsage which does not cease to let new streams of contributions emerge.

---

[36] @Strack Azar, "Dancing in the Street // Silent Music Video," January 13, 2016, YouTube video, 3:00, https://www.youtube.com/watch?v=BHkhIjG0DKc.

3.2 Everyday (Self-)Capture and Re-Appropriation: Audiovisual *Objets Trouvés*

When Jonny Shire walked past a house wall on August 31, 2014, an air conditioner caught his ear. Unmistakably, the loud clicking noise indicated a defect, but it was not technical fascination but rather the resulting rhythmic patterns, which evoked entirely isolated and immediate musical associations, that made him decide to record and upload his discovery directly from his iPhone via the "YouTube Capture" app, adding the title "Broken air con that plays a jazz drum solo!!"[37] In a similar fashion, @Mr. King shares his musical experience of a squeaky door in a Chicago parking lot, even moderating the following "performance": "Okay, so this is awesome. This door is gonna do an impression of Miles Davis' 'Bitches Brew.' Stand by."[38] These captures of readymade sound objects are afforded by mobile phone cameras, letting Jonny Shire and @Mr. King – and millions of other users – imbricate their everyday experience in

---

[37] @Jonny Shire, "Broken air con that plays a jazz drum solo!!," August 31, 2014, YouTube video, 1:05, https://www.youtube.com/watch?v=dmKTfro4g6Y.

[38] @Mr. King, "Door Does Impression of Miles Davis," August 16, 2012, YouTube video, 0:33, https://www.youtube.com/watch?v=wwOipTXvNNo.