

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO E INFORMAÇÃO

MARCELO BERGAMIN CONTER

**IMAGEM-MÚSICA EM VÍDEOS PARA *WEB***

Porto Alegre, fevereiro de 2012.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO E INFORMAÇÃO

Marcelo Bergamin Conter

### **Imagem-Música em Vídeos para *Web***

Dissertação apresentada à banca examinadora do Programa de Pós-Graduação em Comunicação e Informação da Universidade Federal do Rio Grande do Sul (PPGCOM-UFRGS) como requisito parcial para obtenção do título de mestre.

Orientador: Prof. Dr. Alexandre Rocha da Silva

Porto Alegre, fevereiro de 2012.

Marcelo Bergamin Conter

## **Imagem-Música em Vídeos para *Web***

Dissertação apresentada à banca examinadora do Programa de Pós-Graduação em Comunicação e Informação da Universidade Federal do Rio Grande do Sul (PPGCOM-UFRGS) como requisito parcial para obtenção do título de mestre.

Orientador: Prof. Dr. Alexandre Rocha da Silva

Aprovado em: 28 de fevereiro de 2012.

### BANCA EXAMINADORA

---

Profa. Dra. Adriana Amaral

---

Profa. Dra. Ana Paula Penkala

---

Profa. Dra. Nísia Martins do Rosário

---

Profa. Dra. Lizete Dias de Oliveira (suplente)

## AGRADECIMENTOS

Foram várias as pessoas que colaboraram, das mais variadas formas, para que esta dissertação pudesse ser composta. Agradeço Gabriel Saikoski, que nunca hesitou em questionar as hipóteses que eu ensaiava no início do mestrado, além de termos feito leituras em conjunto de alguns dos livros aqui citados; minha colega Susan Liesenberg, com quem enfrentei o preconceito acadêmico aos nossos objetos de análise (“quer dizer que vocês estudam as bobagens da internet?”); meu colega Márcio Telles, que não se envergonha de assumir sua breguice (a ponto de executar *Sexy Yemanjá* em seu notebook durante um GT da Intercom): seu modo de ver a música pop sem dúvida está aqui refletido; André Araújo, pelos debates ofíciosos na portaria da Fabico; as turmas de Seminário de Audiovisual 2011/1 e 2011/2, com as quais não apenas revisei alguns autores aqui utilizados, mas também aprendi muito sobre o objeto deste estudo; à Profa. Suzana Kilpp, com quem fiz minha iniciação científica; meus pais, Bruno e Sandra; e a “excelentíssima”, Giordana Besen. Agradeço também o CNPq.

## RESUMO

A presente dissertação propõe-se a investigar os modos com que a música sobrecodifica a linguagem audiovisual em vídeos para a *web*, criando novos processos de significação e complexificando a virtualidade musical (compreendida aqui como a totalidade irrepresentável de imagens que a expressam). O *site YouTube* se estabelece como um *locus* privilegiado para tal estudo, pois nele encontra-se uma quantidade significativa de vídeos musicais em que a música se manifesta em todos os elementos audiovisuais, interferindo nos processos de composição audiovisual (tanto na trilha visual quanto na sonora). O que pode deste encontro derivar são atualizações e potencialidades do virtual da música, manifestadas como imagem-música. Os vídeos que constituem o *corpus* são organizados em quatro categorias mais recorrentes, ordenados de acordo com a proximidade que têm com a linguagem do videoclipe, o que permite perceber como, progressivamente, nos vídeos para *web*, está ocorrendo um processo de autonomização da imagem-música: *mashup* audiovisual; *sampling* audiovisual; *spoof* de *shreds*; *auto-tunning*. Essas práticas já estavam contidas, em potência, em audiovisuais anteriores, no cinema, na televisão, na videoarte, no videoclipe – e também nas práticas da música eletrônica –, mas somente na *web* elas conseguiram se manifestar a pleno. Como referencial teórico para compor o modo como será observado o fenômeno, utilizar-se-ão as teorias do filósofo Henri Bergson para compreender a virtualidade da música; de Gilles Deleuze, para entender os processos de significação que a sobrecodificação da música exerce sobre o audiovisual, e como ela movimenta suas estruturas através de tal processo; de Nicklas Luhmann, para compor um ponto de observação para os vídeos musicais para *web* diferente do que tradicionalmente se dirige ao videoclipe televisivo; e de Vilém Flusser, para compreender a natureza das imagens técnicas. Conclui-se que os vídeos para *web* estudados efetivamente se apartam da lógica do videoclipe televisivo, parte por não dialogarem com a lógica da indústria fonográfica, parte porque a música é resultado da montagem, e não o contrário, como ocorre normalmente em videoclipes; além disso, evidencia-se que a música é capaz de imprimir algo de si nos audiovisuais estudados, e que está ocorrendo uma tendência para a diluição da distinção entre arte figurativa e música: através desses vídeos, mostra-se possível a produção de música imaginativa, derivada da reciclagem de material audiovisual disponível na *web*.

**Palavras-chave:** Imagem-música. Videoclipe. Virtual. Música. Comunicação. Audiovisual.

## ABSTRACT

The present dissertation proposes to investigate how music over-codifies the audiovisual language of web videos, creating new signification processes and complexifying the musical virtuality (here understood as the unrepresentable totality of images that express it). YouTube's website is established as a privileged locus for such study because it contains a significant amount of music videos in which music manifests itself in every audiovisual element, interfering in the audiovisual composing process (both in the visual and the sound tracks), which is edited according to musical rules. What may come from this encounter are updates and potentialities from the music's virtual, manifested in the form of what will be here understood as image-music. The videos that compose the work's *corpus* are organized in four categories that are defined according to videoclip's language proximity: audiovisual mash up, audiovisual sampling, shreds spoofs, and auto-tuning. This procedure allows us to perceive how, progressively, in web videos, an empowerment of music-image is occurring. These practices were already contained, potentially, in previous audiovisuals, like cinema, television, video art and video clips – and also on electronic music practices –, but only in the web they managed to be fully manifested. Different theories will be presented and utilized as theoretical references, like those of French philosopher Henri Bergson, which will be used to comprehend music's virtuality; works from Gilles Deleuze, will be utilized to aid our understanding of the signification processes that music's over-codification performs on audiovisual, and how its structure is altered by it through this process; from Nicklas Luhmann's theories will provide a different point of view, on web music videos; and Vilém Flusser's works will be employed to understand the nature of the technical image. We conclude that the web videos analyzed effectively move away from the logic of television music video, first, because they do not dialogue with the recording industry's logic; second, because the songs created result from editing and not the other way around, as often occurs in video clips; third, it becomes clear that music is able to print something out of itself on the studied videos, there is an ongoing tendency for the dilution of the distinction between figurative art and music; throughout these videos, the production of imaginative music becomes possible, created from the recycling of audiovisual material available on the web.

**Keywords:** Image-music. Music Video. Virtual. Music. Communication. Audiovisual.

## LISTA DE FIGURAS

Frame 1.....	23
Frame 2.....	23
Frame 3.....	26
Frame 4.....	26
Figura 5.....	64
Frame 6.....	64
Frame 7.....	66
Frame 8.....	66
Frame 9.....	92
Frame 10.....	92
Frame 11.....	92
Frame 12.....	94
Frame 13.....	94
Frame 14.....	94
Frame 15.....	94
Frame 16.....	95
Frame 17.....	96
Frame 18.....	97
Frame 19.....	97
Frame 20.....	97
Frame 21.....	97
Frame 22.....	97
Frame 23.....	97
Frame 24.....	97
Frame 25.....	97
Frame 26.....	101
Frame 27.....	101
Frame 28.....	102
Frame 29.....	102
Frame 30.....	102
Frame 31.....	108
Frame 32.....	109
Frame 33.....	109
Frame 34.....	109
Frame 35.....	109
Frame 36.....	114
Frame 37.....	121
Frame 38.....	121
Frame 39.....	121
Frame 40.....	121
Frame 41.....	121
Frame 42.....	121

## LISTA DE GRÁFICOS

Gráfico 1 .....	48
Gráfico 2 .....	48
Gráfico 3 .....	49
Gráfico 4 .....	98
Gráfico 5 .....	98



## SUMÁRIO

<b>1 INTRODUÇÃO.....</b>	<b>9</b>
<b>2 MÚSICA E AUDIOVISUAL.....</b>	<b>22</b>
2.1 <i>YOUTUBE</i> : PARÓDIA, RECICLAGEM E REMIXAGEM MUSICAL.....	22
2.2 A MÚSICA NO UNIVERSO DAS IMAGENS TÉCNICAS.....	29
2.3 RELAÇÕES ENTRE IMAGEM E MÚSICA NO AUDIOVISUAL: DIFERENTES COMPREENSÕES.....	37
2.4 IMAGEM-MÚSICA.....	44
<b>3 A IMAGEM NO AUDIOVISUAL.....</b>	<b>50</b>
3.1 DEFINIÇÕES DE IMAGEM.....	52
3.2 A IMAGEM SONORA.....	60
<b>4 PERCURSOS METODOLÓGICOS PARA O ESTUDO DA IMAGEM-MÚSICA NOS VÍDEOS PARA A <i>WEB</i>.....</b>	<b>68</b>
4.1 O VIRTUAL DA MÚSICA NOS VÍDEOS PARA <i>WEB</i> .....	68
4.2 A MÚSICA COMO SISTEMA.....	74
4.3 DO NÃO-SENTIDO AO SENTIDO ARTICULADO.....	82
<b>5 A IMAGEM-MUSICA NOS VÍDEOS PARA <i>WEB</i>.....</b>	<b>89</b>
5.1 MASHUPS AUDIOVISUAIS.....	90
5.2 SAMPLING AUDIOVISUAL.....	95
5.3 SPOOFS DE SHREDS.....	106
5.4 AUTO-TUNNING.....	113
5.5 4'33".....	118
<b>6 CONSIDERAÇÕES FINAIS.....</b>	<b>128</b>
REFERÊNCIAS BIBLIOGRÁFICAS.....	131
GLOSSÁRIO.....	136



## 1 INTRODUÇÃO

O *YouTube*<sup>1</sup>, principal *website* repositório de vídeos da internet, conta com uma quantidade enorme de vídeos, das mais variadas categorias. Dentre elas, há uma parcela significativa de vídeos musicais, em que o modo como se apresenta a música pode ser bastante variado: podem ser videoclipes, nos mesmos moldes dos que assistimos na televisão; paródias destes videoclipes, enviadas pelos usuários; vídeos que os usuários publicam, em cuja trilha visual há apenas a imagem da capa de um disco, enquanto que na trilha sonora ouve-se uma de suas faixas; pode ser até a exibição na trilha visual de um disco de vinil rodando numa vitrola, enquanto ouve-se uma canção antiga. Há ainda uma parcela de vídeos que difere, em várias instâncias, do modo como são compostos tradicionalmente videoclipes, e que são justamente os que interessam para esta dissertação: por serem produzidos pelos usuários do *YouTube*, os modos de composição não precisam se alinhar às estratégias de mercado da indústria fonográfica. Tais vídeos apresentam práticas tão diferentes daquelas a que estamos acostumados a ver na MTV que, talvez, não seja possível sequer entendê-los como videoclipes: eles produzem novos modos de relacionar música e audiovisual. Diferentes também das paródias ou do ato de exibir um disco rodando em uma vitrola, o modo como são montados, tanto na trilha sonora quanto na visual, é carregado de uma musicalidade, como se a linguagem musical estivesse se manifestando em todos os elementos do audiovisual. Na maior parte das vezes, trata-se de algum tipo de reutilização de material previamente publicado no *YouTube* e que, muitas vezes, sequer era musical, como reportagens, documentários, entrevistas, seriados, sermões religiosos e comerciais de TV. A montagem, que transforma estes vídeos em música, pode acontecer através da mixagem de dois diferentes vídeos; ou do fracionamento de imagens audiovisuais, como se fossem notas musicais, reordenados para produzir ritmos, versos, melodias e refrões; ou da redublagem de performances ao vivo; ou até da interferência da altura melódica ou do movimento aparente das imagens visuais, através de plugins de softwares de edição de vídeo e áudio.

Desde as primeiras observações deste fenômeno, constatei que os usuários do *YouTube* – que praticaram algum desses tipos de montagem de alguma forma – estavam encarando os softwares de edição audiovisual como se fossem instrumentos musicais tecnológicos. Duas hipóteses surgiram a partir dessa observação. A primeira: a de que o audiovisual é capaz de

---

<sup>1</sup> <http://www.youtube.com>

*imaginar*<sup>2</sup> música, isto é, de fazer com que a música se manifeste não só na trilha sonora, mas também na trilha visual, na relação entre essas trilhas ou ainda através de outros modos (como, por exemplo, na memória do espectador). A segunda: a de que a música, por sua vez, capaz de interferir na montagem a ponto de sobrecodificar o audiovisual e suas diferentes manifestações expressivas. Tais premissas diferem um pouco do modo como estamos habituados a compreender a música no audiovisual. Faz-se necessário introduzir algumas questões sobre nossa percepção e os suportes fonográficos e videográficos.

Quando ouvimos uma gravação musical sendo reproduzida por alto falantes, a fonte que faz as caixas acústicas vibrarem – independente de o suporte ser vinil, CD, MP3 – provém de um registro fonográfico. *Fono*: som; *gráfico*: imagem. Os suportes em si não são música, nem contêm música, mas, ao serem lidos pelos aparelhos de reprodução, as informações neles contidas se convertem em pulsos eletromagnéticos que fazem vibrar as caixas acústicas. Tais vibrações produzem ondas sonoras que são música manifestada sonoramente. Isso quer dizer que, se os suportes não contêm música realizada, mas podem realizar ao decodificarem dados neles contidos, há, nesses dados, música em potência: uma imagem musical.

Assim como os registros fonográficos, nossa percepção também converte o fluxo incessante de ondas sonoras em imagens. Depois de termos ouvido a gravação musical, conseguimos ir, voltar e passar descontinuadamente por todos os trechos da música em nossa memória, e também é possível, se nos concentrarmos, reproduzi-la mentalmente numa velocidade bem próxima de como quando a ouvimos, prova disso é que mais tarde conseguimos cantarolar a canção.

Antes de cantarolar, temos, em nossa memória, imagens da canção, que não são som, nem música, porque não resultam em vibrações de ar. Mas elas potencialmente são som e música, porque podemos tentar reproduzi-las cantando ou tocando um instrumento musical, assim como as imagens sonoras técnicas, que também não são som, mas possuem código e aparelho interpretador de código para converter imagens em sons. Nos dois casos, imagens são elementos de memória: do disco rígido ou das nossas mentes.

Outro aspecto interessante de se notar é que tanto nosso corpo quanto os equipamentos

---

<sup>2</sup>O conceito de imaginação aqui adotado é o mesmo de Vilém Flusser: “[...] significa a capacidade de concretizar o abstrato, e que tal capacidade é novaque [sic]; que foi apenas com a invenção de aparelhos produtores de tecno-imagens que adquirimos tal capacidade; que as gerações anteriores não podiam sequer imaginar o que o termo 'imaginar' significa; que estamos vivendo em mundo imaginário, no mundo das fotografias, dos filmes, do vídeo, de hologramas, mundo radicalmente inimaginável para as gerações precedentes; que esta nossa imaginação ao quadrado ('imaginação<sup>2</sup>'), essa nossa capacidade de olhar o universo pontual de distância superficial a fim de torná-lo concreto, é emergência de nível de consciência novo.” (FLUSSER, 2008, p. 41-42)

de registro e edição fonográficos podem interferir nas imagens que captaram da música. Podem cruzá-las com outras imagens e devolvê-las ao fluxo já modificadas, ou também podem apenas imaginar o que fazer com tais imagens: desenvolver, partindo de imagens potenciais, outras imagens potenciais.

Se nossa percepção e os registros fonográficos são capazes de imaginar música, eles também devem ser capazes de expandir as possibilidades do que possa vir a ser considerado música. Esse é o papel do compositor, que pode tentar fazê-lo através de instrumentos musicais ou manipulando registros fonográficos. Será que é possível também fazê-lo através do audiovisual? Estamos acostumados, ao assistir filmes, à ideia de que a música serve ao cinema, predominantemente, para ambientar narrativas ou expressar determinados estados de espírito dos personagens. Já o que ocorre nos videoclipes, na maior parte das vezes, é complementar com imagens visuais uma canção pré-concebida. Mas nesses dois cenários, fazendo essa leitura, não parece possível pensar na emergência de novas imagens potenciais da música, porque esta parece já ter sido realizada como narração, performance, dança etc. Ao assistirmos um videoclipe, sabemos que a música que ouvimos na trilha sonora já foi pré-gravada. Poderíamos argumentar que os músicos e bandas que aparecem na trilha visual fazendo *playback* estão imaginando a música, mas não podem manifestá-la, pois não são capazes de modificar a canção realizada na trilha sonora. O que eles conseguem imaginar e modificar é a sua performance, que se manifesta apenas visualmente a partir da música.

Talvez não seja possível encontrar novas possibilidades para a Música nesses casos citados porque o modo como observamos o audiovisual tende a ser o de estudar uma música realizada, manifestada na trilha sonora, e que se articula com as imagens visuais.

Por outro lado, para os propósitos desta pesquisa, a Música<sup>3</sup>, em sua *virtualidade* (BERGSON, 2005), é concebida como a totalidade irrepresentável daquilo que possa vir a ser considerado música, expressa sonoramente na trilha sonora audiovisual, mas também como imagens musicais, que podem se manifestar de outros modos, até mesmo visualmente.

O virtual da Música está sempre em constante mutação, pois ele se expande a cada nova imagem musical criada (como, por exemplo, um novo estilo musical), o que é compreendido como uma atualização do seu virtual. “[...] é próprio da virtualidade existir de tal modo que ela se atualize ao diferenciar-se e que seja forçada a atualizar-se, a criar linhas de diferenciação para atualizar-se” (DELEUZE, 2004, p. 78). Essas imagens podem ser de duas ordens: como uma imagem atual, um acontecimento *sui generis*, que reconfigura o

<sup>3</sup>Daqui em diante, quando Música for escrita com maiúscula, estarei me referindo ao seu virtual, e com minúscula, a uma música qualquer, manifestada sonoramente. Pede-se a atenção do leitor para tal diferenciação, que é fundamental para o entendimento da hipótese deste estudo.

virtual, expandindo-lhe as possibilidades de realização; ou como uma imagem possível, que pode ou não se realizar. Para compreender o que é o Virtual, é preciso compreender também o que é o Atual, o Possível e o Real, o que Pierre Lévy (2005) denomina O Quadrívio Ontológico. Ao contrário de como o senso comum propõe, a oposição Virtual/Real é falsa:

[...] o virtual não se opõe ao real, mas sim ao atual. Contrariamente ao possível, estático e já constituído, o virtual é como o complexo problemático, o nó de tendências ou de forças que acompanha uma situação, um acontecimento, um objeto ou uma entidade qualquer, e que chama um processo de resolução: a atualização. Esse complexo problemático pertence à entidade considerada e constitui inclusive uma de suas dimensões maiores. (LÉVY, 2005, p. 16)

Assim, ao invés de a música<sup>4</sup> ser um elemento que trabalha pelo audiovisual, em conjunto com a fotografia, narrativa, montagem, sua virtualidade passa a problematizar o audiovisual, a sobrecodificar sua linguagem. O que pode emergir de tal processo? Novas manifestações sonoras? Provavelmente, mas, mais do que isto, sem dúvida o que pode deste encontro derivar são novas potencialidades para a Música, ou seja, uma problematização de seu virtual.

Posteriormente, com a possibilidade de publicar vídeos na *internet*, em *sites* como o *YouTube*, um novo cenário se desenvolveu para a Música se relacionar com o audiovisual. Diferente da televisão e do cinema, os vídeos publicados em *sites* repositórios de vídeo estão alinhados com fundamentos da cibercultura, como a democratização da produção (todo usuário pode produzir e publicar vídeos), a pirataria e a destituição do conceito de autoria. As três mudanças a seguir parecem as principais para a constituição desse cenário:

- Popularização de *softwares* de edição não-linear de vídeo: a maior parte dos computadores vem de fábrica, há pelo menos oito anos, com Windows Movie Maker (nos PCs) ou com iMovie (nos Macs). Isso colaborou para uma democratização da produção audiovisual ainda maior que a proporcionada pelas câmeras de vídeo, pois tais *softwares* permitiram que se pudesse inserir nas trilhas de vídeo e de áudio qualquer tipo de arquivo;

- Popularização de *softwares* de edição de áudio multipista<sup>5</sup>: para PCs, foram desenvolvidos programas como Cubase e ProTools. A Apple, por sua vez, vende o Macintosh com o GarageBand incluso. Esses programas, antes restritos para produção de áudio, passaram a ser utilizados na remixagem de trilhas sonoras de vídeos, sendo usados em cooperação com os softwares editores de vídeo;

<sup>4</sup>Como um atual, por isso em minúsculas.

<sup>5</sup>São programas especialmente desenvolvidos para gravar canções. Ao invés de gravar todos os músicos em simultâneo, estes softwares permitem que se grave uma trilha de bateria, por exemplo, e, em seguida, a partir da audição simultânea desta, gravar os outros instrumentos, um por um.

- Lógica de lucratividade de *sites* de hospedagem de vídeo: ao invés dos vídeos divulgarem indiretamente um produto (como um disco), a publicidade está imbricada em vários elementos do *site*, seja em *banners*, seja inserindo publicidade dentro da janela do vídeo. Então, o *YouTube* não interfere no conteúdo do vídeo, e procura os *trends* (vídeos que recebem muitas visualizações em um curto espaço de tempo) para inserir publicidade. Uma porcentagem vai para o dono do canal em que o vídeo está hospedado.

Uma vez que as funções de “copiar” e “colar” são parte da base da experiência do usuário na informática, era de se esperar que a mesma prática fosse utilizada largamente em vídeos publicados na *web*. O material empregado para montagem nos *softwares* de edição audiovisual pelos usuários pode ser imagens capturadas por eles mesmos, ou, o que tem ocorrido com muita frequência, vídeos que já estavam publicados no *YouTube*, que são reciclados e ressignificados ao serem remixados pelos usuários. Enquanto o videoclipe e o cinema usam músicas que já haviam sido criadas, ou que são criadas especificamente para serem usadas no audiovisual, nos vídeos musicais para *web* a música só aparece realizada depois da montagem do vídeo em *software* de edição não-linear. Esta é a grande mudança aqui estudada, e ela demonstra outra questão relevante: se o usuário imaginou uma música (em potência, portanto), e pretende fazê-la se realizar através do audiovisual, o *software* se tornará uma espécie de instrumento musical audiovisual. Além disso, como ele tentará fazer música com um aparelho que não foi desenvolvido para tal finalidade, o audiovisual apresentará limitações; este lidará com a ideia de ser sobrecodificado pela Música nos seus *termos*. Sem dúvida que desse encontro se produzirão novas imagens, tanto para a Música quanto para o audiovisual.

Tal proposta não ocorre tão somente através de sonoridades, mas através de uma outra ordem de entendimento do que seja, ou pode vir a ser, Música. Através da montagem de vários elementos não-reconhecidamente (mas *potencialmente*) musicais, os vídeos musicais para *web* fazem emergir o que nesse projeto se entenderá por *imagem-música*. O conceito foi inicialmente desenvolvido em artigo homônimo por Alexandre Rocha da Silva (2006). Inspirado por Deleuze (1985; 2009) e seus conceitos de imagem-movimento e imagem-tempo, e pelas considerações de Nöth e Santaella sobre computação gráfica e música (2009), o autor encontrou, nas performances televisivas de Elis Regina na extinta TV Excelsior, manifestações potenciais daquilo que somente a linguagem computacional poderia ser capaz de fazer: produzir, a partir de um mesmo código (o binário), imagens visuais e sons musicais. Os gestos de Elis, por sua vez, forçaram, em meados da década de 1960, a televisão a imaginar a música (que no caso era, muito pertinentemente, *Arrastão*), pois os operadores de

câmeras eram forçados a produzir *zoom in* e *out* várias vezes durante a execução da peça, prática que não era habitual à época. Nisto Silva reconheceu um ensaio daquilo que hoje o computador é capaz de fazer: imaginar músicas, musicalizar imagens. Um texto sincrético – em que importam mais as articulações entre o visual e o musical do que seus significados independentes – que abre caminho para a emergência de imagem-música.

Nos vídeos para *web*, podemos reconhecer duas ordens de imagem-música: como uma *atualização* do virtual da Música, ao se manifestar através do audiovisual, coordenando os sentidos a serem por ele produzidos, e como a expansão das *possibilidades* desse virtual. Trata-se de imagens que, ao emergirem do encontro da Música com o audiovisual, reconfiguram a virtualidade da Música, expandindo-a e lhe possibilitando novos modos de manifestação.

O desafio aqui proposto é o de reconhecer imagem-música nos vídeos para *web*, e para realizar tal propósito parto de algumas reflexões de Henri Bergson. Para o autor, somos capazes de capturar instantes (atuais), mas o que ocorre *em fluxo*, entre um instante e outro, nossa percepção não dá conta. Assim, Bergson distingue estes dois tempos. Aquele que não podemos apreender, somente intuir e experimentar, e que está em fluxo, é virtual – o seu modo de ser. Todos os instantâneos captados dessa transição, todo congelamento da matéria, é um atual – o seu modo de agir. O virtual gera novos atuais seus, que, por sua vez, o reconfiguram incessantemente – é a *Duração*, e tudo dele faz parte.

Todas as coisas são um misto de virtual e atual, que se modificam constantemente pela ação do tempo. Para Bergson, tudo está em movimento, tudo *é* movimento. “Portanto, quer se trate do dentro ou do fora, de nós ou das coisas, a realidade é a própria mobilidade. Era o que eu expressava ao dizer que há mudança, mas não há coisas que mudam” (2006c, p. 17). Um dos procedimentos de análise que faço na dissertação é o de reconhecer atualizações de imagem-música nos vídeos para *web* a partir das teorias bergsonianas. Ao contrário da virtualidade da Música, que é inapreensível, seus atuais aparecem como uma espécie de figura, de espacialização do virtual. Mas só os reconheço porque primeiro estabeleci o que queria compreender como *virtual*. Para Bergson, a abordagem que devemos ter ao observar fenômenos é a de partir do entendimento do virtual para então reconhecer seus atuais. É a partir do todo que se reconhecerão as partes.

Gilles Deleuze contribui aqui especificamente para a produção de sentidos através da linguagem audiovisual e musical, bem como das estruturas e dos processos de diferenciação e repetição que ocorrem nos mais diversos elementos dos vídeos. Deleuze (2007) entende que o sentido emerge da relação entre elementos desprovidos de significado, e que quando se



produz um sentido novo, criação de formas, é capaz de movimentar as estruturas, como do encontro sui generis de um elemento Musical com um audiovisual, que faz expandir a Música, dando-lhe novas possibilidades de atualização. A repetição de tais acontecimentos (que pode ser a repetição de um refrão em uma música, por exemplo), por sua vez, também tem sua importância na sedimentação de elementos estruturais ou na territorialização, de acordo com o livro *Mil Platôs* (1997), que Deleuze escreveu com Félix Guattari. Diferença e repetição deixam de ser meras oposições e passam a ser repensados em suas funções e ações. A pertinência de seus textos, nessa dissertação, reside na necessidade de compreender como a imagem-música resultante dos vídeos aqui estudados problematiza o virtual da Música e se compõe de diferenças e repetições de elementos audiovisuais (isso ocorre, basicamente, através dos procedimentos de copiar e colar, o que será abordado nos capítulos 2 e 5). Uma vez que abordarei tempo e audiovisual, os tomos sobre cinema de Deleuze, *A imagem-movimento* (1985) e *A imagem-tempo* (2009) também influenciam diretamente na minha observação empírica e também nas análises dos vídeos musicais para *web*.

Nicklas Luhmann (2009) vai colaborar especialmente para a construção do ponto de observação que quero propor, que é o de, partindo da Música, compreender como esta pode sobrecodificar processos de significação no audiovisual. O autor nos ensina a repensar a ideia de observação, trocando a distinção sujeito/objeto por meio/sistema. Vendo os objetos agora como sistemas, interessará entender como eles se adaptam às mudanças do meio em que estão inseridos. No que concerne a este estudo, tanto Música quanto audiovisual podem ser considerados meio ou sistema, dependendo do ponto de observação que se queira tomar. Duas leituras serão feitas: primeiro entendendo a Música como sistema e o audiovisual como meio, para poder compreender como a primeira é capaz de se diferenciar de si ao se atualizar através dos vídeos para *web* (um modo de compreender algo da virtualidade da Música); depois, entendendo a Música como meio e o audiovisual como sistema, sendo este por aquela irritado (um modo de reconhecer os atuais e possíveis da Música que emergem desse encontro). Para se adaptar às irritações sofridas pelo meio, o sistema precisa abrir frestas para observar o meio no qual está inserido, processo que Luhmann denomina de acoplamento estrutural. Através destas “frestas”, o sistema é irritado pelo meio, adquirindo informações novas, que serão por ele trabalhadas de modo a aumentar sua complexidade.

Vale perceber que há pontos comuns entre os conceitos de virtual em Bergson, de estrutura em Deleuze, e de Sistema em Luhmann. Os três, em consonância, permitem o desenvolvimento de um percurso metodológico em direção à virtualidade da Música e suas potências que se atualizam com mais frequência nos vídeos para *web* que constituirão o

corpus.

São utilizadas também as teorias de Vilém Flusser para compreender a natureza das imagens técnicas nos vídeos para *web*. Para o autor, estamos vivendo na *Pós-História*, pois as superfícies dotadas de imagens se impuseram em nossa cultura de tal forma que nos lançaram para fora daquilo que compreendemos por *História*: a descrição dos acontecimentos através de um código linear. Para Flusser, ao nos projetarmos para fora da História, estaríamos dando início a um pensamento em superfície, através de imagens técnicas. A forma como nos comunicaremos nesta Pós-História também é pós-alfabética. “Mas o que agora nos diz respeito existencialmente é o penoso salto do linear [a escrita, que é o modo como a História se desenvolveu e existe] ao adimensional [...], ao 'quântico', ao sintetizável [...]. A exigência que se nos coloca é a de ousarmos dar o salto na nova imaginação” (FLUSSER, 2007, p. 176). A nova imaginação, a que Flusser se refere, é produzida por imagens técnicas. Essas estão um nível acima das imagens tradicionais, como a pintura, pois não mais representam o mundo, mas conceitos do mundo. Ultrapassar a imagem aparente, o caráter mágico de tais imagens, penetrando no âmago da técnica, compreendendo tais imagens não como visões de mundo, mas como produtoras de conceitos do mundo, permitirá que possamos compreender parte deste novo paradigma em que estaríamos inseridos, segundo Flusser. Sem dúvida que os vídeos aqui abordados são compostos exclusivamente de tais imagens, e que eles apresentam uma linguagem de superfície, diferente da linear (Histórica), e que precisa ser problematizada e estudada.

Quando for necessário responder questões musicais, recorrer-se-á aos textos de Wisnik (1999) e Schafer (1991; 2001) para a apresentação das bases do sistema tonal, a reflexão de aspectos técnicos, como o advento de tecnologias de gravação e reprodução sonora (e como elas afetaram a Música) e a reflexão de questões históricas, por exemplo, refletindo em quais aspectos a Música sofreu com as alterações da paisagem sonora durante o início da modernidade. São itens importantes para as análises a serem feitas nesta dissertação, pois assim poder-se-á compreender com mais clareza o cenário técnico e histórico no qual os vídeos musicais para a *web* se constituem, além de possibilitar a compreensão de códigos musicais e sonoros.

Por fim, John Cage também exerce um papel importante na ultrapassagem dos códigos linguísticos que nos bloqueiam de ouvir os sons simplesmente como sons. Trata-se de um dos compositores mais importantes da contemporaneidade, fundamentalmente por suas obras indeterminadas – nelas, o que reina é a imprevisibilidade – não há repetições de sequências melódicas, por exemplo, e sim uma tendência para a pura diferença. Cage pretendia libertar o

ouvinte de uma escuta que procura primeiro as notas, ou, melhor dizendo, o sentido dos sons, abrindo a percepção para uma escuta sem sentido, fora da linguagem. De sua obra, interessa aqui, em especial, a problematização da linguagem musical, que nos possibilita melhor compreender os procedimentos composicionais dos vídeos musicais para *web*.

Falando em procedimentos de composição, cabe neste momento explicitar como foi selecionado o corpus de análise. Não é possível fazer um levantamento quantitativo, por exemplo, porque o recorte que proponho implica em reconhecer subjetivamente as qualidades dos vídeos; é preciso assisti-los; e, em alguns casos, isso deve ser realizado várias vezes para compreendê-los. Não adiantaria, portanto, filtrar por palavras-chave na busca do *YouTube*, por exemplo. O caminho adotado foi o da pesquisa exploratória, não só dentro do próprio *YouTube*, navegando pelos vídeos relacionados e pela busca, mas também através vídeos compartilhados em redes sociais como o *Twitter*<sup>6</sup> e o *Facebook*<sup>7</sup>; de vídeos sugeridos por amigos; e, por fim, em blogs e *websites* especializados, como é o caso de *Know Your Meme*<sup>8</sup>. Assisti a muitos vídeos musicais, não só durante os dois anos que esta pesquisa foi desenvolvida, mas desde 2006, quando o *YouTube* começou a se tornar popular no Brasil, e fui aos poucos reconhecendo que alguns desses vídeos apresentavam práticas diferentes do que se compreende por videoclipe. Embora a quantidade de vídeos citados nesta dissertação está em torno de quarenta exemplos, isso está muito longe da quantidade de vídeos que constam no *YouTube* e que têm as mesmas características. A abordagem aqui proposta, no entanto, é a de encontrar práticas recorrentes e analisá-las em vídeos mais significativos.

Durante a pesquisa exploratória, pude classificar os vídeos que produziam alguma ordem de diferença na Música em quatro categorias mais recorrentes. Apresento-as ao leitor a seguir, e estão organizadas de acordo com a proximidade que têm com a linguagem do videoclipe, o que permite perceber como, progressivamente, nos vídeos para *web*, está ocorrendo um processo de autonomização da imagem-música. Todos derivam de práticas da música popular e eletrônica.

Uma das primeiras práticas encontradas no *YouTube* trata de articular uma música, uma canção popular, na trilha sonora, com imagens de arquivo na trilha visual. Ainda que utilize uma canção pré-concebida, o que a difere do videoclipe é a utilização de imagens visuais que, aparentemente, não têm relação alguma com a canção. Da produção brasileira, a maior parte dos vídeos é composta por cenas do seriado *Chaves*, sincronizando as falas dos personagens como se estivessem cantando uma canção de rock. Compreendo tal encontro

---

<sup>6</sup><http://www.twitter.com>

<sup>7</sup><http://www.facebook.com>

<sup>8</sup><http://knowyourmeme.com>

como “*mashup* audiovisual”. Enquanto o *mashup* no mundo da música eletrônica lida com a ideia de mixar duas músicas diferentes, o *mashup* audiovisual trataria de mixar dois audiovisuais diferentes, ou um vídeo e uma música que provêm de origens diferentes.

O segundo tipo mais recorrente imita a noção de *sampling*, muito recorrente na música eletrônica, e que diz respeito à técnica de produzir amostragens sonoras que são repetidas em *loops*. Usuários do *YouTube* começaram a compor vídeos musicais através da amostragem de planos audiovisuais: neste caso, cada plano contém um som musical acompanhado da imagem visual de um músico a executando: cada plano dura o tempo que a nota dura para a canção ser composta; então a noção de plano perde força, podendo talvez ser melhor compreendida como *Sampling* audiovisual. Ocorre aqui um afastamento maior das práticas do videoclipe, porque, neste caso, plano visual e evento sonoro são inseparáveis, e, principalmente, a canção só pôde ser composta através do *software* de edição de vídeo.

O processo do terceiro tipo é mais ou menos o inverso do que ocorre nos *mashups* audiovisuais. Partindo de performances de famosos guitarristas executando solos virtuosísticos ao vivo (esse ato é conhecido na *web* por *shredding*), usuários do *YouTube* fazem paródias cortando o áudio original e o substituindo por uma outra performance sonora, feita por eles mesmos, mimetizando os movimentos, mas os executando muito mal, como se os guitarristas nunca tivessem pego um instrumento musical na mão.

O quarto tipo parece ser o que melhor expressa o modo com que a Música coordena os sentidos do audiovisual. Trechos de noticiários, depoimentos e entrevistas televisivas são remixados de modo a adquirir certa ritmicidade, como nos casos dos *samplers* audiovisuais, e ainda por cima são transformados em melodias através de um *software* chamado Auto-Tune, capaz de converter qualquer evento sonoro em melodia musical.

Por fim, no final da dissertação, estão os vídeos cuja descrição prometem ser uma execução da peça 4'33<sup>9</sup>, de John Cage, que não necessariamente se enquadram nas outras quatro práticas, mas servem para mostrar como o *YouTube* pode estar ainda mais repleto de *imagem-música* do que esta dissertação foi capaz de evidenciar, além de colaborar na compreensão de alguns conceitos apresentados na dissertação, tais como o de diferença e repetição de elementos musicais, as imagens técnicas e o silêncio.

A dissertação está dividida em seis capítulos – sendo este o primeiro. No segundo, início descrevendo como se desenvolveu o cenário no qual os vídeos musicais para *web* se

---

<sup>9</sup>Datada de 1952, a peça, que dura exatamente o tempo proposto em seu título, deve ser executada pelo seu intérprete sem que ele produza qualquer som com seu instrumento, rompendo assim com a distinção sujeito/objeto, uma vez que os ruídos produzidos pelo público passam a fazer parte da paisagem sonora da composição.

encontram e se configuram. Em seguida, apresento o paradigma pós-histórico identificado por Flusser (2011b) no qual os vídeos musicais para *web* se constituem, e como o princípio do “rolo de fita” (das fitas cassetes, da espiral do vinil, mas também da fita cinematográfica) interfere (ou como pode interferir) nos métodos de composição musical e audiovisual. Em seguida, problematizo a forma como se tem estudado audiovisuais musicais, especialmente o videoclipe, explicitando como neles a Música era subordinada à narrativa ou a aspectos visuais, o que permite observar, em comparação, que nos vídeos para *web* que constituem o corpus, a Música passa a interferir no audiovisual, sobrecodificando seus processos de significação. Por fim, apresento o conceito de imagem-música, como expresso em Silva (2006), para então problematizá-lo, atualizando-o para esta dissertação: enquanto, para Silva, a performance de Elis era apenas um ensaio, um devir de imagem-música, aqui ela será compreendida como uma manifestação plena do conceito, porque, como o ponto de vista aqui é o de reconhecer processos de diferenciação da Música através do audiovisual, e, uma vez que isso de fato ocorre na performance de *Arrastão*, estaríamos diante de um audiovisual que promove uma atualização da Música e ao mesmo tempo estabelece diretrizes para novas possibilidades de atualização, fazendo expandir sua virtualidade.

Como o terceiro capítulo já esboça uma leitura fundamentalmente através das imagens produzidas pelo audiovisual (imagem técnica, imagem sonora, imagem-música...), apresento, no capítulo seguinte, algumas definições de imagem que dizem respeito, direta ou indiretamente, àquilo que esta dissertação compreende por imagem, passando pela semiótica, semiologia e formalismo russo (no cinema). Em seguida, apresento os conceitos de *imagem-técnica*, de Vilém Flusser (2007; 2008; 2011a), para compreender os processos de significação e de imaginação que os aparelhos técnicos (*softwares* de edição de vídeo, *plugins*, discos rígidos) promovem através da configuração de imagens técnicas (que podem ser tanto visuais quanto sonoras) e imagem, de Bergson (1990), para quem todo o universo é composto por imagens, o que colabora para que seja possível reconhecer imagens-música provindas não só da trilha visual, mas também da trilha sonora, *entre* essas trilhas, e até mesmo do choque das imagens que o audiovisual promove à percepção de quem assiste aos vídeos musicais para *web* com as imagens-lembrança que este possui em sua memória. Sigo para um breve estudo do desenvolvimento do uso do som em audiovisuais. Avança-se nisso de modo a compreender a imagem como um produto da técnica, mas também compreendê-la como um produto da percepção. Considero os sons nos audiovisuais também capazes de produzir imagens, e reconheço tal capacidade em seu desenvolvimento no cinema, desde a inserção de músicos na sala executando peças durante a projeção, produzindo imagens para a narrativa, até o cinema

moderno, onde o som finalmente se emancipa da trilha visual e passa a produzir imagens em si e para si.

No quarto capítulo desenvolvo uma metodologia específica de análise dos vídeos musicais para *web*: primeiro, compreendendo, a partir de Bergson, o virtual da Música, para, então, reconhecer processos de diferenciação de si nos vídeos musicais para *web*. Em seguida, promovo a transição do ponto de vista em que normalmente se observam audiovisuais musicais, através das teorias de Luhmann, como já foi anteriormente explicitado. Por fim, trato de como o sentido emerge do não-sentido, das unidades mínimas, dos fragmentos de linguagem audiovisual e Musical que, articulados, criam sentidos: muitos dos vídeos a serem abordados se constituem a partir do fracionamento de planos audiovisuais, a ponto de isolar nota musical por nota musical, formando os *samplers* audiovisuais, que se assemelham aos fonemas na língua (não produzem sentido sozinhos) e que, rearticulados, fazem surgir novos significados.

O quinto capítulo encerra com o reconhecimento e análise da imagem-música provocada pela sobrecodificação que a Música exerce sobre a linguagem audiovisual em vídeos para *web*, utilizando as metodologias e teorias estabelecidas pelo capítulo anterior. Evidenciam-se os processos de diferenciação da Música nesses audiovisuais. Os vídeos são divididos nas categorias que apresentei anteriormente: *mashups* audiovisuais; *sampling* audiovisual; *spoofs* de *shreds*; *auto-tunning*; e por fim, vídeos da peça 4'33", de John Cage, um tipo de vídeo que fica um pouco deslocado das práticas dos outros quatro, mas que necessitam serem debatidos aqui para retomar algumas questões que surgiram no desenvolvimento do referencial teórico.

Este estudo está voltado, portanto, para:

- O entendimento das novas práticas de composição musical que se desenvolveram através, fundamentalmente, de *softwares* de edição não-linear de vídeo (em conjunto com *softwares* de produção de áudio, que exercem funções coadjuvantes) publicados no *YouTube*;
- A compreensão das imagens que a Música desenvolve para si através desse processo;
- A problematização do ponto de vista que se adota ao observar as relações entre Música e audiovisual.

A relevância desse estudo se dá pela necessidade da pesquisa em comunicação compreender os modos como Música e audiovisual se relacionam no campo dos vídeos musicais para *web*, pois os métodos de composição, as estratégias de divulgação e os processos de significação, nesses vídeos, vêm se afastando progressivamente das práticas vigentes no cinema e no videoclipe. Mesmo que seja um cenário novo, de pouco mais de seis

anos, a quantidade de vídeos e a força com a qual eles vêm modificando os padrões estéticos Musicais e audiovisuais já respondem à necessidade de uma observação diferenciada para tais objetos.

Mas, mais do que isso, essa pesquisa pretende, através desses vídeos, poder repensar a relação entre Música e audiovisual através de um olhar que privilegia a observação dos processos de diferenciação da Música, colocando-a em posição de destaque nos procedimentos de análise.

## 2 MÚSICA E AUDIOVISUAL

### 2.1 YOUTUBE: PARÓDIA, RECICLAGEM E REMIXAGEM MUSICAL

Quando o *YouTube* começou a se popularizar em meados de 2005, não demorou muito para que surgissem no *site* vídeos que reciclassem outros vídeos. Em 27 de novembro de 2007, o usuário SuperYoshi publicou o vídeo *The Adventures of Super Mario Bros. 3 REMIXED!!!*<sup>10</sup>, onde simplesmente tratou de modificar e brincar com a ordem dos planos e diálogos de um seriado de televisão. O próprio usuário comenta na descrição do vídeo o seguinte: “Este é a mãe de todos esses vídeos estúpidos<sup>11</sup>”. Provavelmente, a quantidade e a variedade desse tipo de remixagem que se encontram no *YouTube* não se devem unicamente a este vídeo, é claro, mas é curioso que entre os usuários se criou uma nomenclatura curiosa para essa prática, intitulada de *YouTube Poop*. Mais tarde, em 21 de outubro de 2007, outro vídeo usando trechos do seriado *Super Mario World* foi publicado pelo usuário Vorhias: *YTPMV: You Are An Egghead*<sup>12</sup>. A especificidade deste reside no fato de o autor ter criado uma música instrumental, com bateria e baixo eletrônicos, e sincronizado os diálogos do seriado com os compassos da música, fazendo com que os personagens cantassem. YTPMV, como o título do vídeo se refere, significa *YouTube Poop Music Video*. E *music video* é como se chamam os videoclipes em países de língua inglesa.

Por esse e muitos outros motivos que se revelarão ao longo desta dissertação, se tornou muito difícil precisar o problema de pesquisa aqui enfrentado sem estabelecer parâmetros de comparação entre a linguagem do videoclipe com a dos vídeos analisados nesta dissertação. Sem dúvida que algo do videoclipe se imprime nos vídeos musicais para *web*, mas a tendência é que estes se projetem para fora dos seus parâmetros, especialmente por não terem vínculos diretos com a indústria fonográfica. É possível reconhecer que até os clipes massivos, aqueles produzidos para lançar em larga escala o disco de um artista, vêm sofrendo interferências dos fenômenos que ocorrem na *web* e que são promovidos pelos usuários, a um tempo espectadores e produtores de imagens técnicas.

---

<sup>10</sup>Vale salientar que o vídeo foi finalizado e publicado originalmente em [www.sheezyart.com](http://www.sheezyart.com) em dezembro de 2004, quando foi aclamado pelos usuários desse endereço. Como demonstrarei mais adiante, à época em que este vídeo foi postado no YouTube (final de 2007), já havia outros vídeos com a mesma lógica publicados no *site*. Disponível em: <<http://www.youtube.com/watch?v=suSSdHr8sog>>. Acesso em: 24 nov. 2011.

<sup>11</sup>*This is the mother of all these stupid videos*. Tradução do autor.

<sup>12</sup>Disponível em: <<http://www.youtube.com/watch?v=txwB-fYpM98>>. Acesso em: 23 nov. 2011.



Antes da popularização da internet, os fãs estavam limitados a dar audiência à MTV, assistir aos shows e comprar discos, camisetas e outros produtos relacionados às bandas. Ficavam muito distantes, portanto, de interferir nas práticas tanto do artista quanto da indústria, limitados a uma opção dual: comprar ou não comprar.

Com a internet, as possibilidades de interferência aumentaram. Isso não quer dizer que qualquer usuário consegue interferir diretamente nas práticas da indústria, mas a produção coletiva deles faz emergir uma espécie de crítica aos produtos midiáticos. Em 2008, Beyoncé Knowles lançou o clipe *Single Ladies* no seu canal oficial no *YouTube*<sup>13</sup>. Em menos de um mês depois do lançamento, dezenas de *spoofs* – paródias do vídeo – foram disponibilizadas por fãs anônimos, que aparecem dançando em seus quartos ao som da canção. Esses vídeo-respostas funcionam tanto para evidenciar um elogio quanto um pensamento dos usuários sobre esses vídeos, brigando por atenção com o vídeo original na página de busca por palavra-chave. Ao buscar por *Single Ladies* no *YouTube* em 10 de janeiro de 2011, na primeira página, o primeiro vídeo na busca era o original, seguido de 22 paródias, covers e *spoofs*. O site acusou aproximadamente 589 mil vídeos postados com estas palavras-chave<sup>14</sup>.

Então aquele usuário que estiver buscando por um clipe específico pode acabar encontrando dezenas de versões. E versões que chamam muita atenção: enquanto o vídeo original (*Frame 1*) contava com aproximadamente 150 milhões de exibições em 23 de novembro de 2011, um dos *spoofs* mais vistos (*Frame 2*)<sup>15</sup> somava aproximadamente 20 milhões de exibições no mesmo dia. A produção audiovisual dos usuários começava a se tornar cada vez mais significativa, inclusive para eles mesmos.



Frame 1



Frame 2

Isso ficou bem evidente quando a banda OK Go postou o vídeo *Here it goes again*<sup>16</sup>,

<sup>13</sup>Disponível em: <<http://www.youtube.com/watch?v=4m1EFMoRFvY>>. Acesso em 23 nov. 2011.

<sup>14</sup>Este novo comportamento de consumo de música popular pelos fãs foi estudado por Thiago Soares (2009).

<sup>15</sup>Disponível em: <<http://www.youtube.com/watch?v=RJIPEHL85Ig>>. Acesso em 23 nov. 2011.

<sup>16</sup>O vídeo foi originalmente postado neste endereço: <<http://www.youtube.com/watch?v=pv5zWaTEVki>> No entanto, depois que a OK Go assinou contrato com a EMI (um das maiores gravadoras da indústria da música),

em julho de 2006, um videoclipe caseiro que apresenta os músicos fazendo malabarismos em oito esteiras de exercício. Em quatro anos, as exibições do vídeo ultrapassaram a marca dos 55 milhões (contagem de agosto de 2010). Dificilmente um vídeo independente obteria uma divulgação dessa ordem na televisão sem passar por uma distribuidora de grande porte. Este fenômeno configurou o *YouTube* como espaço de excelência na *web* para divulgação videográfica de bandas independentes.

Mas esse era só o primeiro passo. Apesar do reconhecimento que a banda obteve, apontada como um fenômeno da *web* mesmo sendo um clipe caseiro, ele não apresenta nada que não lembre um típico videoclipe televisivo e massivo (assim como o clipe de Beyoncé)<sup>17</sup>. E como veremos nos exemplos que seguem, mais do que um espaço para bandas independentes se mostrarem, o *YouTube* permite o surgimento de novas propostas de relações entre audiovisual e Música.

Se antes, na MTV, as superproduções gravadas em película chamavam muito mais atenção, agora no *YouTube* o interesse se dirige ao vídeo feito em casa. Estamos vivenciando um novo estágio, autônomo e paralelo à lógica do mercado. Segundo Costa (2010, p. 3),

É à luz destas três mudanças fundamentais – a proliferação de inúmeros portais de partilha de vídeos em que o vídeo musical assume um incontestável protagonismo; a consequente explosão do número de vídeos musicais produzidos; e o advento de novas tecnologias, tanto ao nível de [sic] *software* como de *hardware*, que impulsionaram a sua produção e disseminação – que é particularmente pertinente instituir o vídeo musical como objecto de estudo dos novos paradigmas de informação e comunicação em plataformas digitais.

Ao longo dos primeiros anos do *YouTube*, não só videoclipes independentes como o de *OK Go* fariam enorme sucesso, como também novas práticas de relacionar Música com audiovisual começam ser desenvolvidas pelos usuários. Ao contrário da lógica dos videoclipes – audiovisual cuja trilha sonora é preenchida por músicas pré-gravadas em estúdios e lançada em álbuns, e a trilha visual, por imagens dos músicos ou de outras coisas que costumam ter alguma relação com a música –, esses vídeos apresentam novos métodos de composição audiovisual. Certa vez, uma menina pastora deu um sermão aos berros e isso foi gravado em vídeo. O usuário JuCarani publicou esse vídeo no *YouTube* em outubro de 2006<sup>18</sup>, e em março do ano seguinte, Raphael Mendes, do *site* Bobagento<sup>19</sup>, publicava o *Funk da*

---

este link foi bloqueado pela gravadora, que não recebia pagamentos por visualizações. A EMI então publicou novamente o vídeo, agora em seu canal oficial: <<http://www.youtube.com/watch?v=dTAAsCNK7RA>> Para mais informações: <<http://www.guardian.co.uk/technology/pda/2010/jan/20/emi-blocks-ok-go-youtube>>

<sup>17</sup>O que talvez pudesse ser contestado pelo caráter “caseiro” que as imagens visuais apresentam, mas lembro que há uma parcela considerável de videoclipes televisivos com tal estética, vide toda a videografia da banda Devo e o clipe *Low* (2004, Jesse Peretz) dos Foo Fighters.

<sup>18</sup>Disponível em: <[http://www.youtube.com/watch?v=KC0on\\_LpoMM](http://www.youtube.com/watch?v=KC0on_LpoMM)> Acesso em: 24 nov. 2011.

<sup>19</sup>Disponível em: <<http://bobagento.com/>> Acesso em: 8 ago. 2011.

*Menina Pastora*<sup>20</sup>, uma montagem com trechos do vídeo *Menina Pastora Louca* sob a batida de um funk carioca. As falas da menina foram recortadas e reordenadas, entrando em sincronia com a música instrumental.

Este é apenas um dos vários exemplos de remixagem que podem ser encontrados na *web*. Para ficar ainda no exemplo anterior, Raphael Mendes fez pelo menos sete vídeos diferentes usando o mesmo método. À época, já havia uma quantidade bastante intrigante desses vídeos feitos por pessoas em diferentes partes do mundo (como exemplificou-se com os *YTPMV*). Neles fui reconhecendo diferentes práticas. Artistas que produziram seus próprios bancos de imagens para compor seus vídeos; ou que mixaram a trilha sonora de um vídeo com a trilha visual de outro; ou que dublam, na trilha sonora, o que aparece na trilha visual; e assim por diante. Os vídeos musicais para *web* que interessam aqui fazem, portanto, um movimento diferente dos videoclipes massivos: através da montagem de vários elementos não-reconhecidamente (mas *potencialmente*) musicais, fazem *emergir* o que nesta dissertação se entenderá por *imagem-música*. Esse conceito será abordado adiante com mais detalhes, mas por enquanto basta dizer que essa imagem-música não será tão somente resultado de uma montagem sonora, mas audiovisual. Fragmentos e descontinuidades aparecem visualmente nesses vídeos, e configuram uma das marcas desta estética, na qual a Música imprime algo de si no sonoro, mas também no visual desses vídeos. É como se quem produz essas obras fosse capaz de encontrar imagem-música em potência em tudo o que vê (e ouve). Na trilha visual, as imagens se organizam de acordo com a remixagem sonora: se uma frase ou melodia for picotada e rearranjada, assim também se dará visualmente. Então, a montagem visual não é de ordem cinematográfica (plano americano, geral, detalhe...); ela vai respeitar as durações das unidades musicais.

Se antes, no cinema clássico e nos videoclipes massivos, a Música era predominantemente subordinada a aspectos visuais como a narrativa, agora ela predominantemente interfere na linguagem audiovisual. O que se quer observar aqui é como ela faz isso.

Para poder pensar nisso, foi preciso encontrar vídeos musicais no *YouTube* que fossem capazes de problematizar essa questão de pesquisa. Como “pedra fundamental”, há o vídeo *Amateur*<sup>21</sup>, publicado no *YouTube* em novembro de 2006. Lasse Gjersten, um então desconhecido sueco de 22 anos, se registrou em vídeo sentado diante de bateria e piano, tocando apenas uma vez e isoladamente peças e teclas de cada instrumento. Em um editor

---

<sup>20</sup>Disponível em: <<http://www.videolog.tv/video.php?id=214328>> Acesso em: 13 mar. 2011.

<sup>21</sup>Disponível em: <<http://www.youtube.com/watch?v=JzqumbhfxRo>> Acesso em: 13 mar. 2011.

eletrônico de vídeo, ele lançou na linha do tempo as batidas da bateria de modo a compor ritmos (*Frame 3*), e depois repetiu o procedimento com o piano, dividindo a tela em duas molduras e mostrando ambos os instrumentos sendo tocados por ele, em simultâneo (*Frame 4*). Cada batida ou nota é um plano, e ao se repetirem, é também a repetição desse mesmo plano. Temos como resultado na trilha sonora uma música com fluxo contínuo, enquanto que, na trilha visual, imagens descontínuas.



*Frame 3*



*Frame 4*

Reconheço nas células audiovisuais que compõem o vídeo um procedimento de montagem similar à produção de música eletrônica, em que a parte instrumental é elaborada através de amostras sonoras, chamadas de *samplers*: sons autônomos que, ao serem colocados em *loop* ou em relação sincrônica ou diacrônica com outros *samplers* são capazes de produzir ritmos, harmonias, melodias; enfim, de produzir música. Quando Gjersten cria amostras que são a um tempo sonoras e visuais, ele está impondo um método de composição musical à composição audiovisual. O instrumento musical aqui não é a bateria ou o piano, mas o *software* de edição de vídeo. As notas, ritmos e células musicais são *samplers* audiovisuais de pequenas frases ou notas musicais produzidas pelos instrumentos, que no *software* aparecem como possibilidades de criação musical. E o produto final é como uma caixinha de música tecnológica, que reproduz sons de acordo com a maneira como foram programados<sup>22</sup>.

Vários vídeos repetem a ideia desenvolvida por Lasse Gjersten em *Amateur* de montar as canções no software de edição de vídeo, sem fazer *playback*. Essas propostas se repetem incessantemente em cada vídeo, e por fim o que parece mudar efetivamente são apenas as canções. “O que o público aprecia nessa espécie de consumo não é, em última instância, o fator 'originalidade', mas sim, talvez, 'a repetição e suas mínimas variações'” (ECO, 1989 apud FELINTO, 2008, p. 39). É comum na *web* a exploração intensa de *memes* (eventos que são

<sup>22</sup>Antes de *Amateur*, Gjersten experimentou esse processo anteriormente em *Hyperactive*, disponível em <<http://www.youtube.com/watch?v=o9698TqtY4A>> onde faz um *beatbox* também a partir de *samplers* audiovisuais. No entanto, *Amateur* responde mais às questões deste estudo.

rapidamente disseminados pela rede). Em 2007, um vídeo em que aparece o personagem Seu Madruga do seriado *Chaves* cantando, mas com o áudio da música *Florentina* do artista brasileiro Tiririca<sup>23</sup> desencadeou não só milhões de exibições, mas estimulou que muitos outros usuários usassem as mesmas imagens para fazer Seu Madruga cantar dezenas de outras canções. E este é só um exemplo; o mesmo ocorreu e continua ocorrendo com vários outros *memes*. Mais recentemente, o videoclipe *Lotus Flower*, da banda Radiohead, em que o vocalista Thom Yorke aparece dançando freneticamente, foi utilizado da mesma forma, resultando em vídeos nos quais o músico aparece dançando ao som de *Single Ladies*<sup>24</sup>, *Dancing Queen*<sup>25</sup> e até o tema de carnaval da Rede Globo<sup>26</sup>.

Essa é uma prática inspirada nas remixagens de música eletrônica: os *mashups*. Neles, os DJs misturam duas músicas diferentes de modo a criar uma terceira. Geralmente a mistura é entre apenas a voz de uma canção e os instrumentos de outra. No caso dos vídeos para *web*, usa-se um vídeo sem áudio e uma canção (que substitui o áudio original, portanto), e as imagens são reeditadas para que possam se sincronizar à nova trilha sonora. Em *System of Vila - Chaves Suey*<sup>27</sup>, a canção *Chop Suey*, da banda System Of a Down, é acompanhada por cenas clássicas do seriado *Chaves*. Estas são sincronizadas de modo a sugerir que os personagens estejam dublando, ao cantar, tocar e dançar no andamento da música. Esse tipo de montagem é um dos mais recorrentes, provavelmente devido à facilidade de obtenção e edição do material.

Mas não há apenas essas duas práticas (de *sampling* e de *mashing up* audiovisual) na *web*. Além destes, há a série de vídeos *Shreds*<sup>28</sup>, do finlandês Santeri Ojala, que substitui o áudio de solos de guitarristas famosos por uma performance propositadamente abaixo do medíocre, mimetizando os movimentos de mão e palhetada dos músicos dos vídeos. É também o caso dos remixes de reportagens de pessoas bêbadas ou com algum jeito diferente de falar, convertidos em música na edição de suas falas para que se sincronizem com um ritmo de funk carioca ou de *hip hop*, que misturam elementos do que estou chamando aqui de *mashup* e *sampler* audiovisuais. E ainda o da série de vídeos *Auto-Tune the News*<sup>29</sup>, em que a

<sup>23</sup> *Seu Madruga cantando Florentina! (Música do Tiririca) - Chaves*. Disponível em: <<http://www.youtube.com/watch?v=SmJ2eBP-IKM>>. Acesso em: 13 mar. 2011.

<sup>24</sup> Thom Yorke: All The Single Ladies dance. Disponível em: <<http://www.youtube.com/watch?v=KujlXtH20l8>>. Acesso em: 13 mar. 2011.

<sup>25</sup> Thom Yorke- Dancing Queen. Disponível em: <<http://www.youtube.com/watch?v=-43geT-qnwI>>. Acesso em: 13 mar. 2011.

<sup>26</sup> Radiohead Lotus Flower - Thom Yorke é GLOBEZE. Disponível em: <[http://www.youtube.com/watch?v=c1WpkT1\\_r-s](http://www.youtube.com/watch?v=c1WpkT1_r-s)>. Acesso em: 13 mar. 2011.

<sup>27</sup> Disponível em: <<http://www.youtube.com/watch?v=umv5YjYmJ30>>. Acesso em: 13 mar. 2011.

<sup>28</sup> <<http://www.stsanders.com>>. Acesso em: 13 mar. 2011.

<sup>29</sup> <<http://www.barelypolitical.com/autotune>>. Acesso em 13 mar. 2011.

voz de apresentadores de jornais televisivos são sintetizadas com um afinador digital, fazendo com que cantem uma melodia que é sincronizada a uma música instrumental.

São estes os tipos de vídeo que interessam a esta pesquisa. O *corpus*, portanto, será composto por vídeos musicais da *web*, selecionados por pesquisa exploratória, especialmente os vídeos em que a relação entre Música e audiovisual se dá de algum modo diferente do habitual, ou em que essa relação é problematizada. Interessante notar que nestes exemplos (e em muitos outros que ilustrarão esta dissertação) a música emerge em decorrência da montagem audiovisual – depois da montagem, e não o que a motiva. Ela é meio *e* fim. É uma prática tão distante da dos videoclipes tipicamente televisivos que talvez nem possamos chamá-los mais assim. É como se o software de edição audiovisual funcionasse como um instrumento musical (a capacidade desse suposto instrumento será abordada nas análises). Seria um espaço onde a montagem deriva da Música, e não da linguagem audiovisual. É a esse propósito que esta pesquisa pretende se arriscar. Pode a Música sobrecodificar sua linguagem à linguagem audiovisual? E pode ela, deste novo paradigma, sair diferenciada de si?

Partindo deste ponto de vista, os videomakers músicos da internet são aqueles que encontram música em tudo. Será necessário encontrar, portanto, outro ponto de vista para observar os vídeos musicais. Uma aproximação de teorias da música com as da comunicação e do audiovisual se fará necessária. É com o cruzamento dessas teorias que imagino ser possível ultrapassar a opacidade dos procedimentos técnicos e descrever e analisar as estéticas que constroem esses novos fenômenos comunicacionais.

## 2.2 A MÚSICA NO UNIVERSO DAS IMAGENS TÉCNICAS

Nas próximas páginas vamos tratar mais especificamente da natureza técnica dos vídeos que compõem o *corpus* desta pesquisa. Para tanto, seguiremos Vilém Flusser, que antecipou a situação atual de criação de vídeos musicais amadores, quando escreveu as seguintes palavras em 1985:

[...] somente os nossos netos sintetizadores de imagens estarão realmente aptos a compor música de câmera com imagens. Para eles, a nossa distinção (schopenhaueriana) entre “arte figurativa” e “música” perderá todo sentido; eles produzirão imagens precisamente por serem compositores. Por isto proponho, em resposta a Schopenhauer, que o universo das tecno-imagens seja considerado o mundo da música imaginativa. (2008, p. 147)

Naquela época, a tecnologia digital para captação e reprodução de registros visuais e sonoros ainda era uma promessa para os usuários. O CD chegava timidamente em alguns lares, e o aparelho de videocassete, a câmera de vídeo, o toca fitas e o gravador de som eram as únicas opções de gravação, edição e reprodução caseira de vídeos e sons. Mas Flusser já percebia uma tendência para a liquefação das fronteiras entre o audiovisual e o musical nestes aparelhos.

Diferentes das imagens técnicas do cinema e da televisão, as dos vídeos para *web* estão em um nível tecnológico ainda mais complexo. Os vídeos musicais selecionados para análise só puderam se desenvolver como tais e/ou obter reconhecimento na rede devido a vários aspectos inovadores que a internet proporcionou para a Música e para o audiovisual, que, aliás, muito tem a ver com outra previsão de Flusser, quando escreve sobre como seria uma sociedade dominada pelas imagens técnicas. Conforme Hartmann (apud CHAGAS, 2008, p. 2),

As formas de vida deste futuro já são visíveis no nosso ambiente; elas apontam para duas tendências básicas e divergentes: por um lado, a perspectiva negativa de uma sociedade totalitária de *receptores de imagens*, por outro lado, o projeto positivo de uma sociedade telemática de *criadores de imagens*. A primeira opção configura o cenário do tecnofascismo [sic] que reduz os humanos à condição de seres programados; o segundo exorta-nos a explorar o potencial libertador das imagens técnicas e a possibilidade de programar a nossa própria realidade.

Atualmente, convivemos com ambos os cenários. É claro que não são extremos opostos, mas concomitantes, e a situação de receptor não nos levou a um tecnofascismo (ao menos, não ainda!). De qualquer modo, nos importa aqui a possibilidade de qualquer usuário do aparelho poder, se quiser, se transformar em produtor de imagens. Em se tratando do

objeto de pesquisa, pode-se notar um processo cumulativo. Primeiro, os usuários podiam apenas adquirir ou ser receptores das imagens técnicas (cinema, televisão); em seguida, podiam, além disso, registrar suas próprias imagens ou copiar imagens para si (câmeras de vídeo, videocassete); finalmente, hoje, podem produzir e compartilhar com outros usuários as suas imagens ou até mesmo imagens com direitos autorais (em *sites* repositórios de vídeo, como o *YouTube*).

Desde o fonógrafo, o armazenamento e reprodução de sons passaram por incontáveis suportes. Os mais populares hoje são os formatos digitais, especialmente o MP3, que consegue reduzir o tamanho do arquivo para apenas 10% do seu tamanho original. No final da década de noventa, *softwares* gratuitos que convertiam as canções de CDs em MP3 se popularizavam na mesma época que a internet discada. Não demorou muito para que surgissem os primeiros programas que somavam ambos. O Napster é o caso mais famoso, e foi o que causou a fúria das indústrias fonográficas. Graças a ele, os internautas poderiam ter acesso remoto à biblioteca de MP3 de qualquer outro usuário do mundo, de graça, e sem pagar *royalties* aos músicos, nem à gravadora. Embora tenham conseguido derrubar o Napster, nunca foi possível controlar os downloads ilegais. Da mesma forma, com a banda larga foi possível piratear filmes e videoclipes. Mas este não chega a ser exatamente um problema atual. Desde que a indústria passou a fornecer fitas VHS e K7 virgens, os usuários aprenderam a gravar, editar, organizar e, muitas vezes, compartilhar e vender fitas com conteúdo protegido por direitos autorais. A ideia de compartilhamento de registros sonoros e audiovisuais, portanto, não é nenhuma novidade. O que mudou em especial foi a intensidade e a agilidade com que isso agora acontece.

Em 2005 surge o *YouTube*, na mesma época em que a banda larga começa a se popularizar. No *site*, qualquer usuário pode enviar seus próprios vídeos para que o mundo inteiro os assista, bastando abrir uma conta gratuitamente. Como é difícil controlar o que cada usuário publica, o *site* também é uma gigantesca biblioteca de canções, videoclipes, curtas-metragens e até mesmo filmes (partidos em uma dezena de partes, porque o tempo limite de duração é de quinze minutos para cada vídeo) publicados por usuários que não pagaram direitos autorais para tal uso.

Em simultâneo, o processamento dos computadores ficou rápido o suficiente para que softwares de edição de áudio e de vídeo pudessem ser utilizados por usuários amadores. O *Windows Movie Maker*, que passou a ser distribuído como componente integrante do pacote do Windows em sua versão XP, de 2003, era apresentado como uma alternativa para editar os vídeos pessoais e armazená-los em formato compactado (MPEG). À época, o preço das placas



de vídeo FireWire (desenvolvida para transmissão de arquivos de vídeo digital em tempo real) também chegaram a valores populares e, assim, com uma câmera de vídeo com entrada FireWire era possível editar todo o material no computador.

Na década passada, a capacidade dos discos de armazenamento também não parou de crescer: em 2000, o notebook iBook da Apple era vendido com discos rígidos de 10 gigabytes<sup>30</sup>. Atualmente, o MacBookPro (que substituiu a linha iBook e é mais barato à venda) contém disco rígido de 500 gigabytes<sup>31</sup>. Isto possibilitou a armazenagem e a edição dos vídeos capturados diretamente nos computadores. Vale lembrar que toda essa tecnologia serviu do mesmo modo a músicos: “tornou-se acessível do ponto de vista financeiro e os compositores passaram a trabalhar em suas casas ou estúdios pessoais com seus próprios aparelhos, cujas funções são controladas cada vez mais por processos automáticos” (CHAGAS, 2008, p. 8).

Além disso, um diferencial dos softwares de edição não-linear é que eles permitem a importação de quase todo tipo de registro visual e sonoro. Não é de se surpreender, portanto, a infinidade de vídeos caseiros que de vez em quando nos deparamos no *YouTube* constituídos de fotos de família ou de viagens acompanhados de músicas populares. São os álbuns de fotos digitais e Mp3 que os usuários mantêm arquivados em seus computadores.

Do outro lado, o *YouTube* possui esquemas para evitar que os seus vídeos sejam baixados. Mas os usuários sempre descobrem métodos para burlar o sistema e permitir que se faça o download do vídeo. Provavelmente o uso mais frequente desta prática seja o de republicar um mesmo vídeo que esteja recebendo muitas visualizações. De tal prática decorre outra: a de interferir nesses vídeos através de softwares de edição não-linear. Realizar este tipo de interferência, do ponto de vista legal, não estava previsto nem no programa do software, nem no *YouTube*. O usuário joga contra o esquema de proteção autoral do *YouTube*, que não permite o *download* de seus vídeos, assim como joga contra o software de edição não-linear de vídeo, quando o utiliza como instrumento musical, e não como um programa para fazer ficções, documentários ou *slideshows* de fotos de formatura. Em contrapartida, é preciso notar a ambiguidade dessa situação, pois o mesmo aparelho que restringe legalmente a reprodução ilegal de audiovisuais com direitos autorais é o que permite que o usuário faça tais montagens. As grandes empresas não mais são as que produzem as imagens, mas as que armazenam as imagens produzidas pelos consumidores e as fazem circular. “Não mais *quem possui* tem poder, mas sim *quem programa* informações e as distribui” (FLUSSER, 2011a, p. 71).

<sup>30</sup>Fonte: Wikipédia. Disponível em: <<http://en.wikipedia.org/wiki/IBook>>. Acesso em: 2 jan. 2012.

<sup>31</sup>Fonte: www.apple.com. Disponível em: <<http://www.apple.com/macbookpro/specs-13inch.html>>. Acesso em: 2 jan. 2012.

Democratização da produção de imagens técnicas. Isso só foi possível porque estamos todos conectados, e podemos trocar e transmitir uns para os outros tais imagens. Toda a internet gira em torno e depende delas para seu funcionamento. Esse é o grande aparelho que nos conectaria através de imagens técnicas que Flusser previu. Se compreendermos este fenômeno de acordo com o autor, portanto, ambos os movimentos produzidos pelo usuário (o de baixar vídeos ilegalmente e o de usar os softwares de vídeo para remixá-los) são ações que estavam já previstas pelo aparelho.

Retomando a conversa com Flusser do início deste capítulo, “As imagens técnicas são 'arte pura', no sentido em que apenas a música é arte pura. Eis a razão por que sugiro que, com a emergência do universo das tecno-imagens, adquiri novo nível de consciência, nível no qual se tornou possível fazer música imaginativa” (2008, p. 146). Ora, se a música é arte pura no sentido de que está isenta de ter que criar representações do mundo, há um ponto comum com as imagens técnicas, porque estas também são produtos da imaginação. Visam modificar nossos conceitos sobre o mundo. A diferença reside em que não criticamos o grau de realismo de uma música, enquanto fazemos isso com as imagens técnicas. “O observador confia nas imagens técnicas tanto quanto confia em seus próprios olhos. Quando critica as imagens técnicas (se é que as critica), não o faz enquanto imagens, mas enquanto visões do mundo” (2011a, p. 30). Quando vemos uma reportagem de campo na televisão, encaramos como um evento, como algo que aconteceu *de verdade*, embora qualquer imagem técnica, assim como algo que aconteceu, converte eventos em cenas. Por outro lado, quando ouvimos uma música, não conseguimos encará-la como um relato ou uma descrição documental do mundo; ela é 'arte pura' como diz Flusser. Coloca-nos em um tempo de fantasia.

Para exemplificar isso há o vídeo *MC Jeremias*<sup>32</sup>, assinado por “Viniciux” e postado no *YouTube* em 2006. O autor aproveitou dois vídeos publicados na rede um ano antes<sup>33</sup>, ambos extraídos do programa de TV *Sem Meias Palavras*, exibido pela TVI, de Pernambuco, afiliada do SBT. Neles, um rapaz chamado Jeremias foi levado à delegacia duas vezes por ter dirigido uma motocicleta embriagado e sem carteira. Nas duas ocorrências ele é entrevistado pelo repórter Givanildo Silveira. No primeiro vídeo, Jeremias responde às perguntas do repórter aos gritos e palavrões. O repórter se aproveita da situação e consegue fazer com que Jeremias cante uma música, ainda que ininteligível.

Com um *software* de edição de vídeo, Viniciux produziu uma batida eletrônica e sobre ela (em uma segunda trilha de áudio) sincronizou as falas de Jeremias de modo que se

<sup>32</sup>Disponível em: <<http://www.youtube.com/watch?v=IQb3JiL6gmY>>. Acesso em: 15 dez. 2011.

<sup>33</sup>O primeiro está disponível em <<http://www.youtube.com/watch?v=MsdnufGPG1A>>. Acesso em: 15 dez. 2011. E o segundo, em <<http://www.youtube.com/watch?v=0qdWhUqP6DU>>. Acesso em: 15 dez. 2011.

assemelhasse a um funk carioca. A construção provavelmente só tenha se elucidado na mente do realizador por dois motivos: primeiro, o rapaz embriagado canta durante a entrevista e, segundo, seus gritos se aproximam do modo como os funkeiros cantam, sem ter uma preocupação com afinação ou melodia. Quanto às imagens, “Viniciux” as monta, na batida criada por ele, exatamente como elas se encontravam montadas ao áudio original.

Enquanto as imagens faziam parte do telejornal, eram percebidas como janelas para o mundo. No entanto, é preciso lembrar que “Imagens são códigos que traduzem eventos em situações, processos em cenas” (FLUSSER, 2011a, p. 23). Elas não representam aquilo que as câmeras capturam, mas sim compõem imagens que imaginam o mundo.

Na edição ainda entram as legendas, logotipos do programa e outras imagens adicionadas e sobrepostas no panorama, que ajudam a passar uma sensação de fluidez da entrevista, disfarçando a montagem. Esta reportagem é composta de maneira a organizar imagens descontínuas de maneira que desenvolvam uma narrativa, fazendo com que pareçam contínuas.

Quando as imagens são atualizadas em *MC Jeremias*, elas deixam de fazer parte de uma narrativa, de um esquema sensorio-motor, e passam a apresentar situações óticas e sonoras puras ao invés de uma imagem indireta do tempo (DELEUZE, 1985, 2009). Esse seria o tempo de fantasia, falsificante, que a música e vídeos musicais criam: em geral, não têm o propósito de representar o “real”; pelo contrário, eles imaginam *outro* mundo: um mundo fantástico. Elas apresentam o falso como potência do falso. Nesse momento, Deleuze e Flusser dialogam:

Querer saber se tal o qual programa televisionado é “live”, ou vídeo, se tal personagem na tela é político ou ator representando político passa a ser pergunta “vazia”. Diante de todo programa se põe, não a questão: “qual a sua realidade?”, mas a questão: “como funciona?”. A questão da verdade e falsidade passa a ser a questão do efeito produzido. O que interessa, nos programas, não é o *input*, mas o *output*. (FLUSSER, 2011b, p. 127-128)

A conversão de vídeos não-musicais em musicais através da edição destitui seu realismo mágico e nos faz lembrar do caráter conceitual, imaginativo das imagens técnicas que os compõem. É como se elas nos fizessem (re)conhecer que um tempo de vídeo é também um tempo de fantasia. Estão de acordo com a ideia de Flusser, de que elas visam modificar nossos conceitos do mundo mais do que apresentar visões dele. Elas nos provocam a refletir na possibilidade de que há Música em potência em qualquer fenômeno comunicacional. Se não houver, ao menos temos a segurança de dizer que em qualquer imagem técnica há. E isto é assim porque sabemos, com ou sem muita consciência, que imagens técnicas audiovisuais

ou musicais são montadas usando o princípio do rolo de fita, cujo processo se assemelha ao de composição musical em vários aspectos. Rolos que acompanharam a humanidade nos últimos 120 anos, no cinema; no vídeo cassete; no toca fitas; em gravadores de mão; em filmadoras; em câmeras de vídeo; nos primeiros computadores.

O gravador de fita foi o principal aparelho de reprodução de som da era analógica. Entretanto, ele foi também frequentemente utilizado para sintetizar sons. Por exemplo, repetindo-se um som em loop e criando-se um circuito recursivo entre os sistemas de reprodução e gravação do gravador, pode-se gerar estruturas rítmicas e sons contínuos a partir de impulsos sonoros. (CHAGAS, 2008, p. 10)

Com a fita foi possível criar, dentre vários efeitos, *loops*, efeitos de *feedback*, aceleração ou modificação da duração e inversão (como o solo de guitarra em *Tomorrow Never Knows* dos Beatles, do álbum *Revolver*). Interessante notar também que nas partituras podemos encontrar algumas marcações que depois seriam imitadas pelas técnicas de edição das fitas. Existe notação para indicar quando um trecho deve ser repetido da mesma forma, como quando se copia e cola adiante um trecho de fita; existem as inversões, quando uma melodia é repetida, mais adiante na música, mas de trás para frente; há os ‘acelerandos’ e ‘ritardandos’; e assim por diante.

Nos primórdios do cinema, o projetista divertia o público rodando o filme mais rápido ou devagar, ou até de trás para frente. O mesmo pode-se fazer com fitas de vídeo e de áudio. Aqueles que conviveram com o toca fitas provavelmente tiveram de lidar alguma vez com um aparelho desregulado que lia o lado inverso da fita, tocando as músicas de trás para frente. Brincadeiras como as de apertar simultaneamente o botão *play* e *fast forward* nos apresentaram sons fantásticos, velozes e agudos. O vídeo cassete também nos apresentou essa possibilidade com as imagens visuais, além do congelamento (*pause still*).

A gravação de música em disco ou fita afetou a composição. Todos os sistemas ordenados de linguagem requerem redundância. A música é um desses sistemas, e sua redundância consiste na repetição e recapitulação do material principal. Quando Mozart repetia um tema seis ou oito vezes, era para ajudar a memória a guardá-lo, para recuperá-lo mais tarde. [...] De fato, a função da indústria de gravação de fornecer redundância e, conseqüentemente, estabilidade à vida, numa época em que o futuro parece incerto, não deve ser subestimada, e se o sucesso das estações de rádio que repetem sempre as mesmas canções servir como alguma indicação, os seres humanos não ignoram esse valor. (SCHAFER, 2001, p. 165)

Levanto esses exemplos para mostrar que quando brincávamos com estes aparelhos, também estávamos jogando contra seus programas. Flusser comenta isso em *A filosofia da caixa preta* (2011), quando fala que os programas dos aparelhos podem ser enganados, permitindo a introdução de elementos humanos não-previstos, como os tais efeitos obtidos na

violação do programa “toca fita”. O que acontece em seguida é que a indústria (o aparelho que desenvolve aparelhos) lança novos aparelhos que integram esses elementos que os usuários inventaram ao jogar contra.

Com esse conhecimento acumulado, os produtores de imagens técnicas sentam diante dos softwares de edição não-linear de seus notebooks de última geração para criar os seus vídeos. Quando vão aplicar efeitos, encontram na caixa *de diálogo*, dentre as várias possibilidades, todas aquelas que já haviam praticado brincando com as fitas: *increase speed*; *decrease speed*; *reverse*; *mirror image*; *freeze*; *cut*; *copy*; *duplicate*. Todos estes efeitos de fita foram emulados pelos programas digitais, e esses acabam sendo os métodos mais simples de interferência, pois os produtores já estão predispostos a fazer esse tipo de interferência por serem pré-programados a isso pelas próprias fitas. Estão todos lá, incorporados ao programa. E as possibilidades são ainda maiores, o que proporcionalmente dificulta ainda mais jogar contra o aparelho.

Quando interferimos dessa forma, estamos tratando os audiovisuais menos como representações do mundo e mais como mecanismo de imaginação, palavra que Flusser redefine

no significado aqui pretendido: imaginar é fazer com que aparelhos munidos de teclas computem os elementos pontuais do universo para formarem imagens e destarte, permitirem que vivamos e ajamos concretamente em mundo tornado impalpável, inconcebível e inimaginável por abstração desvairada. [...] Em suma: a definição de “imaginar” foi formulada para articular a revolução epistemológica, ético-política e estética pela qual estamos passando. (2008, p. 45)

Então não é de agora que se recorta, cola, inverte e alteram-se registros audiovisuais e musicais. O que mudou é que aqueles que antes só assistiam a isso agora são os produtores, e que estão dispostos a converter vídeos que antes eram “cassetadas”, jornalísticos ou documentais, em música. E música que não pode ser apenas ouvida, ela é feita também para ser vista – separar a trilha visual da sonora, nestes vídeos, pode desqualificar a montagem evidente.

Esses truques de fita são uma tecnologia datada, muito distante dos efeitos de *super slow motion*, *live 3D* e *Bullet Time* que a grande indústria tem à disposição. Mas interessa pelo uso: um uso que vai contra o aparelho. “O decisivo em relação aos aparelhos não é quem os possui, mas quem esgota o seu programa” (FLUSSER, 2011a, p. 46). Embora capaz, não foi para isso que o *Windows Movie Maker* foi elaborado.

Há uma grande diferença entre um vídeo que apresenta uma música para o que se converte em uma. No primeiro (seu principal representante são os videoclipes massivos), a

música, canção popular, já veio montada e inserida na trilha sonora, enquanto a montagem visual complementa com visuais os sons que ouvimos. Nos vídeos do corpus desta dissertação, a montagem deriva das células musicais, por isso não respeita as normas, as linguagens televisivas e cinematográficas. Ela acontece para que a música aconteça, em todos os seus elementos; a repetição não é mais *flashback*, mas ritmo. O *slow motion* vira ritardando. O *jump cut*, troca de verso pro refrão. Não há sequer linearidade narrativa. A diferenciação entre imagem e som, ou entre imagem e música, começa a ficar nebulosa.

Na imagem técnica, não se trata de *intermix*, mas de mútua superação de música e imagem. Toda música pré-aparelhística e toda imagem pré-aparelhística não passam de elementos de duas tendências convergentes que estão atualmente se juntando. O “mundo da vontade” e o “mundo da representação” convergem. É isto que o termo “audiovisual” procura articular, mas falha, por ser termo proveniente de nível de consciência ultrapassado. (FLUSSER, 2008, p. 146-147)

Audiovisual e a Música foram ambos sendo cada vez mais interferidos pela montagem das fitas. Das antigas pautas musicais, os compositores passaram a compor através de edição de fitas com material gravado, como exemplifica a composição do disco *Bitches Brew* (1970) de Miles Davies. Boa parte do material gravado pelo músico e sua banda foi feita de improviso. Em seguida, e em conjunto com o produtor Teo Macero, Davies finalizou as peças na mesa de edição, modificando as estruturas, cortando trechos dos improvisos ou até adicionando efeitos de *loops* de fita. Ao longo do último século, as práticas composicionais foram tendendo cada vez mais para a manipulação de imagens técnicas pré-gravadas. Os aparelhos de reprodução audiovisual e musical foram tendendo, pouco a pouco, a se transformarem em um mesmo: o computador, no qual a montagem de ambos os materiais ainda simula o processo manual de edição em fita que o precedeu. Mais uma previsão acertada de Flusser, quando, segundo Chagas, diz que

futuramente, as partituras irão desaparecer, substituídas por “fitas de fitas e de fitas” (1996: 176). [...] Assim, a prática musical telemática é organizada como um diálogo cibernético entre “músicos” e “memórias inteligentes” que são, ao mesmo tempo, transmissores e receptores de informação. (2008, p. 7)

Tal experiência evidencia a passagem paradigmática para a pós-história de que fala o autor em *A Filosofia da Caixa Preta* (2011a). Com Flusser e para além dele – porque os vídeos para a *web* aqui estudados não são apenas imagens técnicas, mas experiências estéticas que fazem, a partir de um dado programa, convergir diferentes linguagens igualmente técnicas – ênfase a necessidade de se compreender como tais programas se tornaram a chave para que se identifiquem os desafios do tempo presente. Os vídeos para a *web* aqui apresentados

são a expressão empírica de tal situação em que os antigos instrumentos técnicos são substituídos por aparelhos que os reprogramam a partir das regras que lhes são próprias.

Também começa a evidenciar-se a ideia de que, a partir destes aparelhos mais complexos, a Música torna-se mais capaz de interferir na linguagem audiovisual ao invés de ter uma característica de ambiente, de pano de fundo de uma narrativa, de ocupar apenas a trilha sonora, como tem sido sua relação histórica com o cinema e o videoclipe, o que será explorado a seguir.

### 2.3 RELAÇÕES ENTRE IMAGEM E MÚSICA NO AUDIOVISUAL: DIFERENTES COMPREENSÕES

Gostaria de focar neste momento no videoclipe, pois ele provocou um grande rompimento no modo como as trilhas visual e sonora se relacionavam no cinema e na televisão até seu surgimento; além disso, as principais teorias a seu respeito contribuem para o entendimento da proposta desta dissertação. Início com um levantamento que fiz de alguns autores que abordam este objeto. Na década de 1990, Coelho Netto já enxergava a música na MTV “como constituinte essencial da imagem [...] Na MTV ocorre uma simbiose intensa entre imagem e som musical. Tão intensa que é arriscado apontar a dominante, se a imagem ou o som” (1995, p. 159-160). No mesmo livro, defende que a descontinuidade não era um produto do cinema, mas um traço do pensamento moderno (p. 32). E é essa ideia que Machado destaca no livro “A televisão levada a sério” (2009) como típica do videoclipe. Nos planos de duração curtíssima, misturas de bitolas, suportes e outras estripulias de vanguarda, o autor encontra uma tendência dos videoclipes procurarem uma relação mais plástica e sincrética da parte das imagens. Disso decorre a tendência para a não-linearidade narrativa e a linguagem metonímica, em que o todo se dá através da exibição de pequenas frações suas em planos de duração pequena.

Se opondo a Machado em alguns aspectos, Lillian Coelho (2003) ressalta a importância da narrativa nos cliques, questionando artigos que falavam do não-sentido e da iconicidade como qualidades intrínsecas do videoclipe, quando claramente podia se perceber que a grande maioria deles contam histórias com início, meio e fim, especialmente os de bandas e cantores massivos.

Nessa época, Thiago Soares publica o livro *Videoclipe: o elogio da desarmonia* (2004). Seu entendimento do videoclipe como desarmônico decorre da leitura que o autor faz de Eisenstein, que diz que a montagem cinematográfica provém de um choque entre dois planos, e como consequência desse choque, emerge a desarmonia entre os elementos.

Também com uma produção muito relevante há o norueguês Alf Björnberg (2000), que ainda em 1992 já fazia uma crítica aos estudos sobre videoclipe, especialmente aos autores de maior destaque internacional como Simon Frith (1988) e Andrew Goodwin (1992). Para o autor, a maioria dos estudos da década de 1980 pouco ou nada fala da trilha sonora ou da musicalidade intrínseca dos videoclipes.

A norte-americana Carol Vernallis avança em um outro sentido: em artigo publicado em 2008, estuda a apropriação da linguagem do videoclipe no filme *Brilho eterno de uma mente sem lembranças* (2008), dirigido por Michel Gondry, reconhecido também pelos seus clipes premiados. A autora também escreveu sobre estéticas e o contexto cultural do videoclipe em seu livro *Experiencing Music Video* (2004).

O ponto de vista que vem se adotando para analisar os videoclipes na internet é o da sua migração da tevê para a *web*. Ariane Holzbach e Marildo Narcolini (2009) recentemente abordaram o sucesso de videoclipes de algumas bandas brasileiras como NX Zero na *web*, o que aponta uma transição do público que consome videoclipe da televisão para a internet.

Thiago Soares atualmente também vem estudando o sucesso que fenômenos midiáticos como Beyoncé e Lady Gaga vêm fazendo através da *web*. Segundo ele, “os videoclipes, hoje, não podem ser mais enquadrados como eminentemente televisivos (massivos) ou digitais (pós-massivos). Apostar em tais retrancas significa tirar do clipe a sua característica mais evidente: a constante flutuação por ambientes midiáticos” (2009, p. 2). Parte do sucesso parece se dever aos fanclipes, paródias dos clipes originais feitas pelos fãs e que por sua vez são muito visualizados pelos internautas. Erick Felinto (2008) também estudou as paródias, também conhecidas por *spoofs*, mas para evidenciar uma necessidade de a ciência passar a olhar diferente para aquilo que comumente é entendido por “lixo digital”.

Por fim, encontramos o projeto do português João Pedro da Costa, que como diz o subtítulo de um de seus artigos publicados, apresenta *uma proposta epistemológica para o estudo de vídeos musicais na web 2.0* (2010). Assim como Björnberg, o autor critica autores clássicos, especialmente por terem analisado os videoclipes na maior parte das vezes a partir de teorias literárias. O autor enxerga essa possibilidade no tensionamento de diversos fluxos com os quais o vídeo musical na *web* se envolve: não bastando os do próprio vídeo, entram



aqui também os fluxos de portais de partilha, o navegador, o *hardware*, outras redes sociais e assim por diante.

Quando se lê Arlindo Machado falando da descontinuidade, da linguagem metonímica, Thiago Soares desenvolvendo sua ideia de desarmonia, e Coelho Netto relacionando a MTV com o pensamento moderno, nota-se que estas e outras características designam muito bem o videoclipe como um gênero televisivo. Características estas todas baseadas no que se fez na trilha visual. Mas, enquanto isso, na trilha sonora, tivemos uma repetição incessante de canções com estruturas conservadoras, dotadas de introdução, verso, refrão, re-exposição do verso e do refrão, um solo de guitarra de vez em quando, e encerramento. Aos pesquisadores restaria, neste cenário, interpretar as imagens através de teorias oriundas de outros âmbitos, como bem aponta Costa, fazendo referência a grandes nomes da pesquisa em videoclipes:

a maioria dos autores que se dedicaram ao estudo do vídeo musical nos últimos 25 anos [...] utilizaram, quase sempre de forma inconsciente, ferramentas teóricas e terminologia oriundas dos estudos literários. É o caso da utilização de conceitos como *narrativo* e *performativo* (Andrew Goodwin, Carol Vernallis e Saul Austerlitz); *recepção* (E. Ann Kaplan, Lisa Lewis, John Fiske e Andrew Goodwin); *texto*, *leitura* e *leitor* (John Fiske); *paródia* (Andrew Goodwin); *pastiche* (todos os autores pós-modernistas); *autor* (Jeff Ayeroff, Matt Hanson, Saul Austerlitz e Roger Beebe); *pseudônimo* (Roger Beebe); *close reading*, *flashback* e *elipse* (Carol Vernallis); *mise-en-scène* (Matt Hanson); *cânone* (Saul Austerlitz) e *heteronímia* (Cynthia Fuchs) (2010, p. 5-6).

Não estou negando o potencial explicativo que essas teorias tiveram – e ainda têm – mas preocupo-me aqui com a necessidade de se questionar as qualidades musicais dos audiovisuais a partir de um olhar que também seja musical.

Entendo o videoclipe mais como uma proposta de *inserção* da canção popular dentro da TV. *Ora, mas foi exatamente para isso que ele foi criado!*, poderiam arguir. Mas quando se trata mais de *pensar* a virtualidade da Música do que simplesmente *inserir* uma canção específica, fica difícil encontrar um clipe que opere desse modo, ainda mais em clipes massivos. Não se força a Música a fazer nada porque nunca foi necessário para a indústria, nem para seus espectadores, menos ainda para o audiovisual<sup>34</sup>. Parece um beco sem saída para um pensamento da Música: de um lado, uma indústria que repete velhas fórmulas até que se esgotem; do outro, diretores de videoclipes contratados pela indústria que fazem avançar o estado da arte das imagens videográficas. A Música se estrutura, nesse espaço, no sentido

<sup>34</sup>Alerto que, embora o modo como estou problematizando o videoclipe neste momento possa parecer acusador, não estou fazendo juízo de valor: não há problema algum que determinado clipe não movimente consideravelmente as estruturas da Música. É apenas uma proposta diferente da dos vídeos que constituem o corpus desta pesquisa, que estou aqui diferenciando.

mais duro e estático do termo: ela sedimenta, clipe após clipe, cada vez mais os elementos da canção popular.

Não é à toa que os pesquisadores de videoclipe buscam entender a articulação entre a trilha visual (considerando o que dela é projetado como *imagem*) e a trilha sonora (considerando o que desta é projetado como *música*). Tal dicotomia é produzida de modo a se encontrar, através de relações diferenciais entre um e outro, sentidos, significações; enfim, elementos da ordem da linguagem que possam ser estudados. É uma proposta de se entender o audiovisual como um texto sincrético, cujos sentidos emergem das relações entre imagem e som, mas incorporando a linguagem musical como se fosse *do* audiovisual (parte da trilha sonora). É comum também entender a Música e o audiovisual como linguagens por si. Cada uma, isoladamente, é capaz de produzir suas relações diferenciais: uma pela linha melódica, criada pela diferença de altura entre as notas, e outro pela diferença na troca de enquadramento entre planos, por exemplo; bem como seus elementos simbólicos: a Música com o intervalo de quarta aumentada, por séculos representando o *diabolus in musica*, e o audiovisual usando efeito de eco na voz das personagens para representar um *flashback*; e, como pontos singulares, talvez possamos destacar o modo como cada uma artificializa a espessura do tempo; uma através de *ritardandos*, a outra via câmera lenta.

Dizer que a música não é mais pano de fundo, mas constituinte essencial da imagem como Coelho Netto (1995) diz não é suficiente – além de estar fazendo um elogio apenas às características visuais que decorrem da estética da MTV (porque o autor também compreende a imagem como trilha visual), segue-se ignorando os efeitos que a Música imprime no audiovisual, tanto no visual, quanto no sonoro, quanto na sua virtualidade<sup>35</sup>. A Música exerce, portanto, um papel complementar, coadjuvante da trilha visual e/ou da narrativa. Seja no cinema como música “de fundo”, “de acompanhamento”, “de mistério”, especialmente no cinema *blockbuster*, a Música é funcionária de um motivo maior: a narrativa. Isso acabou por acontecer até nos videoclipes, que surgiram

[...] unicamente como um apoio para a canção. Hoje em dia seu discurso visual acabou dominando a canção de referência e são tão importantes suas imagens como a canção, sendo às vezes muito mais potentes, com linguagem própria e grande força expressiva, com um impacto social enorme. (FERNÁNDEZ SANCHEZ, 1997, p. 158, tradução do autor)<sup>36</sup>

<sup>35</sup>A virtualidade da Música é explorada no capítulo 4.1.

<sup>36</sup>“[...] como únicamente un apoyo para la canción. Hoy día su discurso visual ha terminado dominado a la canción de referencia y son tan importantes sus imágenes como la canción, siendo a veces mucho más potentes, con un lenguaje propio y gran fuerza expresiva, con un impacto social enorme”.

Quando se estudam os videoclipes observando as articulações entre o visual e o sonoro, o que emerge são elementos como as técnicas de corte *on the beat*, a descontinuidade (MACHADO, 2009), a narratividade (COELHO, 2003), a desarmonia (SOARES, 2004), mais para compreender aspectos de códigos culturais do que musicais da pós-modernidade. Na mesma linha de raciocínio, Björnberg ressalta que “a atenção de escritores e acadêmicos pode ser resumida como a quebra de linearidade narrativa, de lógica casual e da coerência temporal e espacial” (2000, p. 348, tradução do autor)<sup>37</sup>. O autor segue apontando a compreensão de diversos autores sobre o videoclipe como sintoma da pós-modernidade. Mas todas essas ideias se prendem aos fenômenos visuais.

Penso que, pela importância que a Música tem para o videoclipe, a atenção que se dá a ela é insuficiente. Björnberg já vinha percebendo essa tendência:

[...] a maior parte dos autores citados tem pouco a dizer a respeito. Propõem que, para o usuário, a música é de alguma forma 'dominada' pelo visual. (Berland 1986; Kinder 1984; Tezlaff 1986) [...]. Kaplan parece considerar os vídeos musicais como baseados principalmente pelas letras das canções (1987: 47 f.), restringindo seu discurso das características musicais para um pouco mais do que meia página (1987: 123 f.) (2000, p. 349-350, tradução do autor)<sup>38</sup>.

Da parte da indústria fonográfica, a tendência nos últimos anos tem sido deixar de lado propostas ousadas em videoclipes. Nos clipes de Lady Gaga, um dos maiores fenômenos midiáticos atualmente, ocorre uma construção conservadora do audiovisual: o clipe *Telephone* (Jonas Åkerlund), de 2009 por exemplo, repete muitos ingredientes da estrutura de *Thriller* (John Landis), de 1983. Em se tratando de videoclipe financiado pela indústria fonográfica, sempre foram e continuam sendo vídeos elaborados a partir de uma música pré-concebida, para a qual se produzirão imagens para se articularem com os elementos musicais, na maior parte das vezes exibindo os músicos que tocam a canção em *playback*.

O videoclipe, tanto para o olhar da indústria fonográfica quanto para a pesquisa acadêmica, parece ter contribuído muito mais para um desenvolvimento do audiovisual do que para a Música<sup>39</sup>. As músicas utilizadas nascem da garagem, do gueto, das *raves*, de conservatórios, de festas, de estúdios. Mas não nascem *dentro de um pensamento* ou *a partir de um* audiovisual. O videoclipe sobrepõe imagens (visuais e sonoras) por cima de uma música entendida como manifestação estritamente sonora. Desse encontro de uma videoarte

<sup>37</sup>“the attention of writers and scholars may be summarized as the breakdown of linear narrativity, of casual logic, and of temporal and spatial coherence.”

<sup>38</sup>“[...] most of the authors cited have remarkably little to say about it. The opinion has been proposed that, for the user, the music is somehow 'dominated' by the visuals. (Berland 1986; Kinder 1984; Tetzlaff; 1986) [...]. Kaplan seems to regard music video visuals as mainly based on song lyrics (1987: 47 f.), restricting [350] her discussion of musical characteristics to little more than one half page (1987: 123 f.)”.

<sup>39</sup>O que não quer dizer que não tenha contribuído de alguma forma à Música.

“à frente do tempo” com músicas com estrutura conservadora<sup>40</sup>, emergem relações que passam a ser mais entre a trilha visual e a trilha sonora do que entre o audiovisual como um todo (seus elementos visuais e sonoros) e a Música. Esta, por sua vez, é tomada de assalto e diminuída no seu poder de diferenciar de si e de provocar uma diferença no audiovisual: sua potência de diferenciação transforma-se em devir minoritário no discurso audiovisual. O videoclipe a absorve para poder fazer avançar as estéticas do audiovisual, o que talvez explique porque, para muitos estudiosos, ele é um gênero *televisivo*.

Nesse contexto eu comecei a desconfiar: não é possível que o videoclipe nasça e desenvolva sua linguagem tomando a Música de assalto para poder se desenvolver, sem devolver a ela também um desenvolvimento de sua linguagem, técnicas e estéticas. Não é possível que a Música, e o rock em especial no caso dos videoclipes, o mesmo que se expandiu tomando também de assalto a linguagem da poesia para constituir as letras de música, a virtuosidade de Paganini para neutralizar e naturalizar solos de guitarra e a construção tonal irresponsável do *blues* não tenha expandido sua virtualidade neste casamento de décadas com o vídeo.

Onde vão se encontrar as linhas de fuga, os procedimentos que poderiam produzir novas figuras para a Música através do audiovisual? Em artigo publicado em 2010, (CONTER; KILPP) na obra de Michel Gondry encontramos um ensaio dessa ordem. No videoclipe *Come into my world* (1996), por exemplo, o diretor chegou a intervir na estrutura da canção para que ela melhor pudesse servir aos propósitos visuais do clipe. *Mas isso é colocar a Música a serviço do audiovisual de novo!*, poder-se-ia arguir, mas o resultado é justamente o oposto: as sincronizações que Gondry faz entre os elementos visuais, os blocos de duração se encaixando com maestria com a duração dos versos, pontes e refrões da canção potencializam suas características, evidenciam a cadência tonal e a repetição típica da música popular na trilha visual. Neste processo, fica bem claro o quanto essas imagens potencializam a estrutura e os elementos da canção que foi escolhida para virar videoclipe. Em outro vídeo, intitulado *Around the world* (1997), o diretor desenvolve uma coreografia composta por quatro grupos de dançarinos. Cada grupo representa visualmente um instrumento musical. O grupo que representa a bateria faz gestos fortes e em *staccato*; o grupo que faz o teclado mimetiza o subir e descer da escala musical subindo e descendo escadas; e assim por diante. Nesse caso, a coreografia ajuda o espectador a perceber isoladamente como e quando soa cada instrumento na música, o que geralmente pode ser uma dificuldade para espectadores que não

---

<sup>40</sup>Parece um dos maiores desafios da canção popular fugir de estruturas compostas por blocos de verso, ponte, refrão e solo.

são músicos. Gondry conseguiu fazer a trilha visual apresentar imagens potencialmente musicais.

Fazendo um brevíssimo ensaio: num ponto de vista que pensasse as relações entre audiovisual e Música, o termo *desarmonia*, que aparece em Soares (2004), talvez não fizesse o menor sentido, pois a canção popular usada nos clipes sempre tende para a *harmonia*: até mesmo por detrás das microfônias do Sonic Youth ouvem-se cadências tonais, temperadas<sup>41</sup> e bem harmônicas que regem sua obra. Nesse mesmo ponto de vista, a ideia de descontinuidade mudaria também. Talvez possa explicar um esforço do vídeo para tentar harmonizar-se com o pulso selvagem do rock, acompanhando sua violência, sua massa sonora que estanca todas as frequências sonoras ao máximo. ”Não é o cinema moderno (ou o audiovisual) que inventa a descontinuidade: isto é um traço do pensamento, da realidade da modernidade” (COELHO NETTO, 1995:32). O rock também é descontínuo e desarmônico: há certo desrespeito às cadências tonais nas estruturas dos acordes – tal que, em muitas canções, cada acorde é independente harmonicamente dos outros. Então, se rock e audiovisual são desarmônicos, ambos estariam entrando em relação de harmonia e de continuidade o tempo todo.

Michel Chion problematiza, em *A audiovisual* (2008), a apropriação de termos musicais pelo audiovisual. Embora seja razoável que a transposição de noções como *leitmotiv* tenha ajudado ao cinema expandir a sua linguagem, essas relações sempre se mostram confusas e arbitrarias:

[...] muitos dos exemplos dados como modelos de contraponto eram, em rigor, casos de harmonia dissonante, uma vez que mostravam apenas uma discordância pontual entre uma imagem e um som relativamente à sua natureza figurativa. Nós próprios, ao utilizarmos a metáfora musical, não devemos ser enganados por ela: o termo harmonia também não descreve a especificidade do fenómeno audiovisual. (p. 35-36)

Mas não quer dizer que só porque os termos musicais foram, historicamente, não “mal apropriados”, mas “tomados de assalto” pelo audiovisual, que a Música não imprima algo de si no audiovisual em videoclipes profissionais, televisivos. É possível que as análises a serem feitas na dissertação, nos vídeos musicais para *web*, possam posteriormente colaborar para esse ponto de vista também em videoclipes e outras formas de expressão audiovisual que utilizam Música. Mas foi da observação empírica dos vídeos musicais para *web* que emergiu o

---

<sup>41</sup>O termo deriva da escala temperada, um padrão de afinação que surge no período barroco, dividindo a escala musical de doze tons matematicamente – e não fisicamente –, o que faz com que alguns intervalos, como o de terça maior (dó e mi, por exemplo) crie um batimento sonoro, portanto desafinado em comparação ao intervalo que as ondas sonoras sugerem. Há desarmonia na Música desde então, quando foi naturalizado um intervalo que não corresponde a fenômenos naturais.

problema de pesquisa, uma intuição que não foi provocada pelo cinema, tampouco pelo videoclipe.

Mas como a Música pode adquirir essa capacidade?

## 2.4 IMAGEM-MÚSICA

Imagem-música é um conceito que foi desenvolvido por Silva (2006) para dar conta de um fenômeno audiovisual específico: “um texto sincrético, quer dizer, um texto formado por diferentes linguagens, em cujos planos de expressão importam mais as articulações entre sons (musicais) e imagens, do que as gramáticas específicas de cada linguagem considerada isoladamente” (p. 41). O autor traz um exemplo extraído da televisão através do qual ele desenvolve o conceito. Ao descrever a performance televisiva de Elis Regina, reconheceu nos seus gestos exagerados (ela ficou conhecida como *Eliscóptero*) uma tentativa de forçar uma montagem audiovisual que se articulasse com a canção que ela canta. Na década de 1960, a linguagem televisiva estava pouco adaptada à exibição de performances ao vivo. Na maior parte do tempo, as imagens visuais apresentam *closes* dos cantores, às vezes sequer exibindo os músicos de apoio. E se exibiam, eram ainda capazes de tirar do quadro os instrumentos, mostrando apenas os rostos. Os gestos de Elis, em contrapartida, faziam com que os operadores de câmeras fossem obrigados a efetuar *zooms ins* e *outs* para alternar os enquadramentos. Um verdadeiro *Arrastão* audiovisual, como propunha o título da canção em questão, faz com que a câmera registre mais a música e menos os músicos. É possível pensar a música (trilha sonora) e a imagem (visual) separadas, mas nesse caso elas são articuladas por Elis para criar um só texto. Este, segundo Silva, seria um devir de imagem-música, pois o conceito só se daria a pleno na informática, capaz de gerar sons e imagens a partir de um mesmo código, o binário. O autor traz a seguinte citação de Nöth e Santaella, que também é relevante para esta dissertação:

[...] o que se tem hoje [...] é uma dissolução de fronteiras entre visualidade e sonoridade, dissolução que se exacerba a um ponto tal que, no universo digital do som e da imagem, não há mais diferenças em seus modos de formar, mas só nos seus modos de aparição, isto é, na maneira como se apresentam para os sentidos (2009, p. 91).

Desse modo, a imagem-música de Silva expressa, conforme consta no resumo de seu texto, “o modo matemático [– derivado do cálculo computacional –] como são produzidos imagens e sons musicais e a variabilidade de suas textualizações” (2006, p. 40). Esse modo é o mesmo das imagens técnicas de Flusser, que, no campo do digital, partem da nulodimensionalidade, pois que as imagens surgem de cálculos (do latim *calculum*, “pedras”, por isso sem dimensão: elas não teriam altura nem largura: são apenas pontos dispersos), que ao serem computados fazem produzir textos (o código binário, unidimensional), que, por sua vez, ao serem interpretados pela máquina, compõem imagem técnica digital (bidimensional). Essa imagem pode representar o mundo, por conter nela uma abstração da terceira dimensão (a profundidade), mas ao mesmo tempo, e mais importante para Flusser, contém disfarçadamente nela os processos anteriormente mencionados e que pretendem esconder-se do resultado final para enganar aquele que admira as imagens técnicas.

Retomando a distinção entre arte figurativa e música imaginativa do capítulo 3.2, quando Flusser se refere à música (no último capítulo do livro *Elogio à superficialidade*, 2008), está se referindo à música erudita de câmara, recorte que circunscreve e diz respeito especificamente àquilo que ele quer tratar: a ideia de que os terminais telemáticos poderiam reunir pessoas de lugares muito distantes pelo intuito de compor música imaginativa. Agora, em Nöth e Santaella, a impressão que dá é que eles estão igualmente se referindo a esse tipo de música, que Wisnik (1999) chama de Tonal, e que teria seu provável “ocaso criativo” no serialismo e no atonalismo. De acordo com Peirce, a tendência da música, *dessa música*<sup>42</sup>, é de se tornar cada vez mais matemática nos seus modos de composição. Na partitura, por exemplo, os intervalos, as durações, os acordes, todos os processos composicionais podem ser calculados.

A matematização da música eletrônica, assim como dos vídeos musicais para *web*, ocorre estritamente no processo nulodimensional, no momento em que são computadas as relações. Torna-se necessário salientar aqui que, enquanto ela é computada, ela ainda não é música manifesta, mas possibilidade de música, algo que será problematizado no terceiro capítulo.

Penso que a montagem que Elis Regina força em *Arrastão* não articula o visual e o sonoro – mas o audiovisual e a Música, pois Elis força a sincretização pensando todos os elementos musicais – a letra, a tonalidade, a textura, até mesmo sua potência visual: enquadramento, fora de quadro, movimentos de câmera, sua postura ao cantar, gesticular... é

<sup>42</sup>Como os vídeos a serem estudados (com exceção dos que são representações de 4'33'') estão inseridos ou no mínimo dialogam com o sistema tonal, não cabe aqui discutir as tendências de outros tipos de música, como a modal.

importante pensar nessa diferença porque um choque entre elementos visuais e sonoros ainda seriam da ordem do audiovisual. É o uso global da Música por Elis, de inserção de musicalidade não só no sonoro, mas principalmente no visual, que produz uma nova ordem no audiovisual. É preciso provocar a diferença, aumentar as possibilidades de acasos na Música.

Por isso, pretendo fazer um desdobramento no conceito: as articulações que fazem a imagem-música não serão então somente entre sons (musicais) e imagens, mas entre a Música e o audiovisual em todos seus elementos. Assim, penso que o conceito possa se prestar à metodologia de análise de audiovisuais musicais. Dessa forma, a Música se emancipa da trilha sonora e torna-se capaz de não só se articular com a linguagem visual do audiovisual, mas também capaz de interferir e criar nessa trilha. O olhar passa para o de buscar imagens musicais, ou melhor, imagem-música, na trilha visual e também na sonora.

A imagem-música, portanto, será, nesta dissertação, mais do que um texto sincrético entre imagem visual e sonora. É um jogo entre aquilo a que a Música deseja atribuir uma musicalidade no audiovisual e este disponibiliza de atributos possíveis para a Música se atualizar. É um jogo entre as duas durações, na forma como elas trabalham uma contra a outra, bem como uma a favor da outra. Onde houver maior tensionamento (ou até mesmo indeterminação) entre as durações da Música e do audiovisual será um lugar privilegiado para a análise a ser realizada por esta pesquisa.

A Música aparece então como produtora de novos problemas (um virtual, portanto), e ao agir no audiovisual, ao se relacionar e jogar com ele, produz as soluções (atualizações: imagem-música).

Nesse movimento de atualização do conceito, o que Silva entendia como um devir de imagem-música em *Arrastão* passa a ser uma imagem-música plena. O devir, ou a potência de imagem-música, por sua vez, passa a ser as potencialidades contidas em seu virtual, conforme será explorado nos capítulos a seguir.

Talvez os gráficos abaixo, baseados na explicação do que é o virtual, de Pierre Lévy (2005), colaborem no entendimento dessa mudança. Lévy faz as distinções a partir de duas dicotomias: virtual e atual, possível e real.

Retomando, sucintamente, os conceitos, o virtual “é como o complexo problemático, o nó de tendências ou de forças que acompanha uma situação, um acontecimento, um objeto ou uma entidade qualquer, e que chama um processo de resolução: a atualização” (LÉVY, 2005, p. 16). Se a totalidade irrepresentável da Música é o seu virtual, ele precisa atualizar suas formas para que continue potente, para que ocorram processos de diferenciação que tornarão esse virtual mais complexo:



A atualização inventa uma solução ao problema colocado pelo virtual. Com isso, não se contenta em reconstruir recursos, nem em colocar uma forma à disposição de um mecanismo de realização. Não: a atualização *inventa uma forma*. Ela cria uma informação radicalmente nova. [...] A temporalidade da atualização é a *dos processos*. (LÉVY, 2005, p. 139)

O possível está relacionado com o virtual (porque ambos são latentes); são elementos que constam como possibilidade de realização. O real, por fim, existe através do movimento incessante de atualização do virtual, provocando a queda de potencial do virtual, mas ao mesmo tempo apresentando novos problemas para este.

No gráfico 1 há setas que apontam os diferentes processos desse quadrívio, organizadas em pares. O primeiro par trata das passagens entre o possível e o real; a primeira seta refere-se ao processo de realização (do possível ao real), que trata da eleição de um possível que deva ser realizado, em detrimento de outros, que se enfraquecem ou se sedimentam no virtual. Lévy (2005, p. 16) alerta que esse processo não é efetivamente criação, pois criar implica na produção de ideia e/ou formas novas, enquanto que a realização é apenas a materialização de uma ideia; a segunda seta faz o caminho inverso da realização. “A potencialização produz ordem e informação, reconstitui os recursos e reservas energéticos” (LÉVY, 2005, p. 138). Cada materialização de possíveis exige do virtual que novas possibilidades sejam compostas, como produção de recursos para a solução de novos problemas. O virtual necessita desse movimento incessante de realização e potencialização para conservar sua existência.

O segundo par de setas refere-se às passagens entre o virtual e o atual. Assim, a terceira seta é a dos processos de atualização, que é quando ocorre uma resolução de um problema contido no virtual. Diferente da materialização de um possível, a atualização é capaz de gerar uma informação *sui generis*, e que, por isso, não estava prevista como possibilidade. Ela é a resolução de tensões contidas no virtual. A última seta é a da virtualização. Logo que ocorre uma atualização, o virtual torna a gerar novos problemas, novas tensões a serem resolvidas.

Os quatro elementos do quadrívio estão diretamente relacionados, e dependem uns dos outros para exercer suas funções.

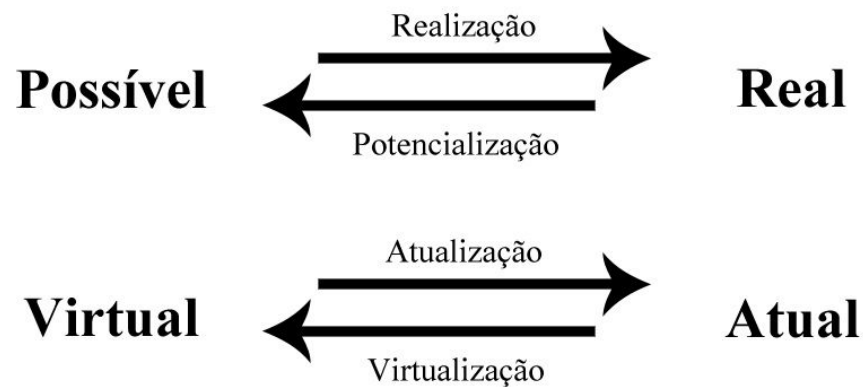


Gráfico 1

Dependendo do que quer ser estudado, diferentes pontos de observação podem ser tomados. A abordagem de Silva, por exemplo, se dá nos termos do gráfico 2. Seu movimento metodológico, representado no gráfico com textos e setas na cor vermelha, parte do entendimento do virtual do audiovisual, as *audiovisualidades*<sup>43</sup>, para então compreender aquilo que se realiza em um audiovisual qualquer, como *Arrastão* cantada por Elis na televisão, por exemplo. Neste audiovisual, opera-se uma distinção entre suas duas trilhas físicas: a visual, que compreende as imagens, e a sonora, que compreende a música. Esse movimento das audiovisualidades ao audiovisual faz produzir novos sincretismos entre o visual e o sonoro do audiovisual (imagem-música), bem como expande o virtual com sincretismos potenciais. Portanto, movimentada as estruturas das audiovisualidades.

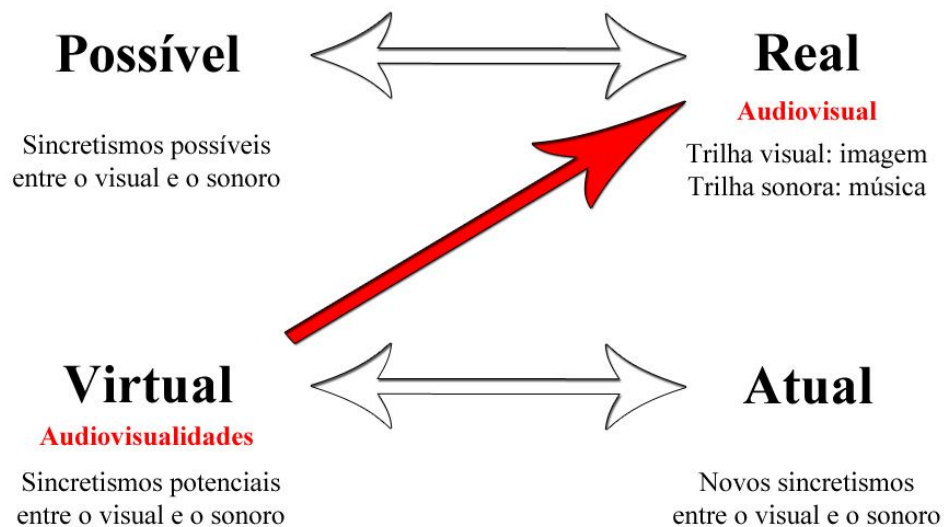


Gráfico 2

<sup>43</sup>O conceito, desenvolvido pelo GPAv (Grupo de Pesquisa Audiovisualidades), trata de compreender as audiovisualidades para além da irreducibilidade das mídias (SILVA; ROSSINI, 2009, p. 7-8).

Essa abordagem é similar a que os pesquisadores de videoclipe tomaram na maior parte das vezes. Ela está inserida em uma área de estudos preocupada com o audiovisual. Por isso necessita-se fazer essa atualização do conceito para fazê-lo produzir para o se quer aqui propor. Como apresentado no gráfico 3, faz-se aqui outro caminho. Metodologicamente, inicia-se os estudos olhando para o objeto, para os vídeos musicais para *web*, e neles já pretende-se observar o audiovisual sendo problematizado pela Música. Parte-se, portanto, do realizado. Nesses vídeos pretende-se encontrar imagens-música, no plural mesmo: interferências da Música no audiovisual, processos de diferenciação de sua virtualidade, resolvidos através de atuais. Reconhecendo as imagens-música atualizadas, pode-se pensar também em imagens-música potenciais, resultantes de novos problemas para o virtual apresentados pelas imagens-música atuais. É nos atuais que se reconhece a sobrecodificação da Música nos audiovisuais estudados, ou das possíveis sobrecodificações, que permanecem latentes. Por fim, deve-se chegar à virtualidade da Música, mais especificamente à parte dessa virtualidade que se manifesta no audiovisual. Está estabelecido, portanto, que este não é um trabalho sobre audiovisuais musicais, mas sobre a interferência da Música no audiovisual.

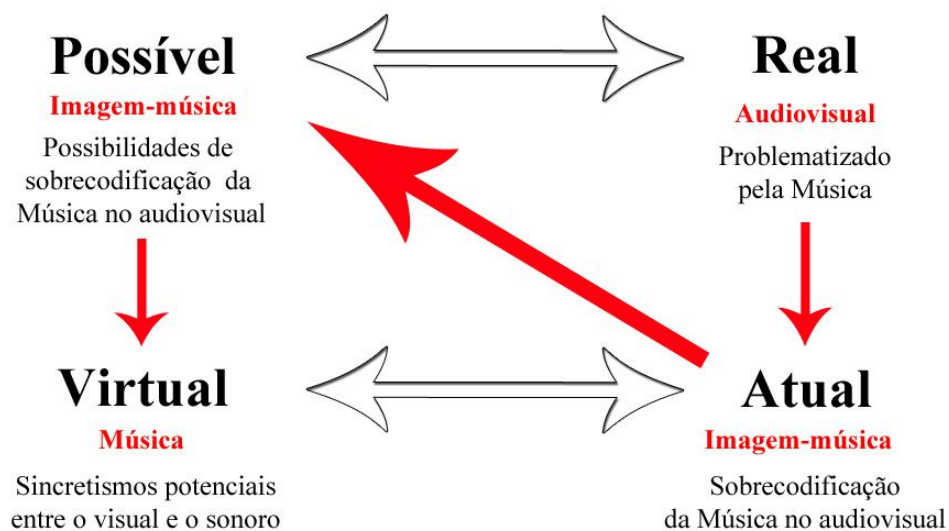


Gráfico 3

É com esse olhar que gostaria de partir para o próximo capítulo, onde iniciarei definindo o que será entendido por *imagem* nesta dissertação, para que fique mais claro onde e em quais elementos dos vídeos musicais para *web* pode se reconhecer imagem-música.

### 3 A IMAGEM NO AUDIOVISUAL

No *Dicionário teórico e crítico de cinema*, Aumont e Marie escrevem as seguintes palavras sobre as trilhas visuais e sonoras:

Tecnicamente, a banda imagem é o suporte pelicular apenas da imagem, não combinada, mas sincronizada, com um som separado (*image track*, em inglês). A banda som é o suporte pelicular unicamente da pista sonora, óptica ou magnética, separada da imagem (*sound track*). (2003, p. 29)

Ao se estudar audiovisual, costuma-se tratar as formas que aparecem nos fotogramas projetados do cinema ou em grupos de *pixels* das telas de vídeo (o visual) como *imagem*, e das vibrações de ar que soam de caixas acústicas (o áudio) como *som*. Esse olhar se dá em qualquer tipo de audiovisual, independente do suporte. Ele é resultado de toda uma tradição científica e filosófica de compreensão de como *vemos* o mundo. Voltemos ao mito da caverna de Platão, mas dessa vez problematizado por Simone Pereira de Sá: a saída que o pensador da caverna encontrou foi através da metáfora da luz, e não do eco, “demarcando uma tradição ocidental de hierarquização dos sentidos liderada pela visão [...]” (2009, p. 108). Ora, esse pensador, que se divertia com as sombras projetadas nas paredes, poderia muito bem se deliciar também com as reverberações das ondas sonoras externas refletidas pelas mesmas paredes. Se houvesse dificuldade de perceber que os ecos não provinham das sombras, não estariam eles *representando* o som do exterior? Não haveria aí a ideia de uma imagem de um som? Flusser já se questionou sobre o sonoro no audiovisual:

Em termos visuais, os filmes são superfícies, mas para o ouvido eles são espaciais. Nadamos no oceano de sons, e ele nos penetra enquanto nos confrontamos com o mundo das imagens, esse mundo que nos circunda. O termo “audiovisual” oculta isso. [...] Essa terceira dimensão, que muda completamente o modo de ler a superfície dos filmes, é um desafio para aqueles que pensam as superfícies, e somente o futuro poderá dizer se isso será resolvido. (2007, p. 109-110)

E Arlindo Machado,

Na verdade, mal entramos ainda numa área nova de estudos que pertence talvez menos ao domínio da musicologia do que a uma possível semiótica das relações audiovisuais. O campo de investigações que se abre nessa área é imensurável. Seria necessário, em primeiro lugar, começar a compreender como funciona a nossa *estereopercepção*. Já conhecemos a *estereoscopia* e a *estereofonia*, que dizem respeito, respectivamente, às sínteses, operadas no cérebro, das informações perceptivas diferenciadas que recebemos através dos dois olhos e dos dois ouvidos. Agora é preciso entender como esses dois pares de informações se combinam também no cérebro para sintetizar uma unidade perceptiva verdadeiramente audiovisual. (2009, p. 165-166)

Como esse capítulo irá demonstrar, estou mais interessado no diálogo com Flusser, em pensar o que normalmente se entende por “som” no audiovisual como algo também capaz de produzir imagens, tanto como suporte, quanto como percebido, isto é, convertido de seu suporte físico (no caso, código binário) em ondas sonoras e interpretado pela nossa percepção; enquanto que, respondendo a Machado, não se tratará aqui de pensar se o visual se impõe ao sonoro, ou o oposto, nem se é uma relação mútua: importa pensar de outro ponto de vista: pensar nas articulações entre audiovisual e Música, o que será mais abordado no próximo capítulo.

Para dar os primeiros passos em direção a um entendimento dos vídeos musicais para *web*, faz-se necessário conceituar especificamente para esta dissertação o termo *imagem*, que aqui diferirá do modo como a pesquisa em comunicação geralmente o problematiza.

Cabe desde já fazer um esclarecimento: nesta dissertação, sempre que for mencionado o termo *trilha visual*, estarei me referindo à exibição visual de informações programadas para serem como tal: a sequência de fotogramas projetado em uma tela no cinema, o panorama videográfico composto por malha de *pixels* em mutação constante da televisão ou, por fim, a moldura que exhibe vídeo digital no panorama videográfico de monitores de computador; e quando mencionar *trilha sonora* estarei falando especificamente das ondas fonográficas que são interpretadas e projetadas por caixas acústicas durante uma exibição audiovisual. A trilha sonora abrange qualquer som produzido por um audiovisual, desde que decodificado de uma informação que tenha sido programada como tal. Fica mais fácil de entender desse modo: no cinema, por exemplo, é trilha sonora aquilo que consta correndo ao lado dos fotogramas, as ondas fonográficas ópticas, e não é trilha sonora o som do projetor ou a inserção de músicos dentro da sala de cinema. Compreendo como *trilha* algo pré-disposto linearmente, que exige que um aparelho interprete-o e converta-o em um composto de luzes ou em um composto de ondas sonoras. Dessa forma, posso trabalhar com o conceito de imagem liberto da trilha visual. Não posso usar como sinônimos *imagem* e *trilha visual*, porque pretende-se aqui reconhecer imagens as mais variadas, não apenas visuais, mas sonoras, mentais, técnicas, musicais...

Mas antes de partir para tal empreitada, faz-se necessário circunscrever o que será aqui entendido por *imagem*. Abordarei a seguir algumas teorias que colaboram para que se chegasse à definição do que será compreendido por imagem (em perspectiva bergsoniana) especificamente para esta dissertação.

### 3.1 DEFINIÇÕES DE IMAGEM

Nöth e Santaella iniciam o primeiro capítulo do livro *Imagem: cognição, semiótica, mídia* dividindo o mundo das imagens em dois domínios: o material, das representações visuais, e o imaterial, das imagens em nossa mente. As mentais remetem a aspectos visuais, que, de acordo com os autores, seriam “visões, fantasias, imaginações, esquemas, modelos, ou, em geral, como representações mentais” (2009, p. 15). O lado perceptível e o lado mental da imagem, ao se reunirem, produziriam o *signo* ou *representação*. Mas, como esta dissertação já vem ensaiando, há outras definições de imagem. Se as imagens materiais só pudessem ser representações visuais, então não poderíamos supor que, assim como as sombras, os ecos que o pensador da caverna de Platão poderia ter ouvido se converteriam, através de sua percepção, em uma imagem mental, como uma representação de um som que está no mundo lá fora.

A questão é que, para Santaella, o som não tem tal capacidade:

O som não tem poder para representar algo que está fora dele. Pode, no máximo, indicar sua própria proveniência, mas não tem capacidade de substituir algo, de estar no lugar de uma outra coisa que não seja ele mesmo. Essa falta de capacidade referencial do som é compensada por seu alto poder de sugestão, o que fundamentalmente o coloca no universo icônico, onde operam as mais puras associações por similaridade [...] (2005, p. 19).

A autora contra-argumentaria, portanto, declarando que o eco que o pensador ouve é o próprio som de fora da caverna: o eco ocorre quando uma onda sonora encontra um obstáculo e muda de direção, mas ainda seria, portanto, a mesma onda. Mas, e se, ao invés de tratarmos de sons “do mundo”, falássemos de sons reproduzidos tecnicamente? Um som de piano que ouvimos de uma gravação em vinil ou MP3, que foi capturado através de microfones, representa sim algo que está fora dele: um piano de verdade, “no mundo”. E um som simulado de piano, produzido por sintetizadores, por exemplo, ainda é capaz de fazer referência a um piano de verdade. Do ponto de vista desta dissertação, os sons – quando reproduzidos tecnicamente, independente de como foram desenvolvidos – possuem sim uma alta capacidade referencial, ainda que seja predominantemente de ordem icônica, como Santaella propõe.

Realmente, a categoria de ícone é concebida por Peirce de forma mais geral e compreende também formas não visuais, por exemplo, acústicas ou mesmo táteis, olfativas ou também formas conceituais de semelhança signica. Há que salientar ainda que, na polissemia do conceito de imagem na linguagem comum, o qual compreende, além das imagens óticas, também imagens acústicas e imagens mentais, existe uma extensão da definição do conceito de imagem que se aproxima daquela extensão do ícone segundo Peirce. (SANTAELLA, 2009, p. 37-38)

Para Charles Sanders Pierce, haveria três classes de signos que poderiam ser aplicados para a compreensão dos mais diversos fenômenos: a primeira são os signos icônicos, que possuem uma relação qualitativa com os objetos que representam. São os signos mais próximos das sensações primeiras que um objeto pode provocar em nosso corpo: quando ouvimos um som<sup>44</sup>, sua iconicidade se dá antes de reconhecermos de que fonte esse som provém, e antes ainda de descobrirmos o que este som significa. Isso pode ser mais fácil de compreender com sons extremamente graves, que se espalham pelo ambiente: pode demorar algum tempo até nos darmos conta, se estamos em um quarto com as janelas fechadas, se um som grave que ouvimos se trata de um caminhão passando na rua, um avião, uma britadeira, um condicionador de ar ou um carro de som (o que nos faz sentir um pouco como o pensador da caverna). Enquanto não reconhecemos o que esse som significa, e de onde ele provém, estamos imersos, predominantemente, na sua iconicidade: na textura desse som, se é alto, baixo, forte, recorrente, circular, instável. Ao reconhecer sua fonte, estamos diante de um signo indicial: sei que é um carro de som, por exemplo. Por fim, pode-se chegar a um terceiro estágio, em que o som que provém desse carro simboliza os problemas de poluição sonora da civilização moderna. Note que esses signos são cumulativos: todo símbolo contém algo de indicial e icônico; todo índice possui uma dada iconicidade. Aplicando a uma reprodução sonora, o som de um piano microfonado, por exemplo, ele é icônico nas qualidades timbrísticas, similar a um piano sendo executado ao vivo; é indicial porque faz referência a este instrumento musical; e poderá ser simbólico dependendo do que os arranjos das notas executadas gerar para o interpretante.

Então poderíamos adicionar à classe de imagens materiais as representações sonoras também. Talvez se pudesse assim desvendar outras formas de imagem material, mas como o objeto de pesquisa aqui proposto é audiovisual, basta-nos como imagens materiais representações visuais e sonoras. Como Nöth e Santaella (2009) já expressaram, tais imagens, sozinhas, não são capazes de significar: elas precisam se encontrar com imagens mentais, aquelas que nossa percepção busca na memória para compreender as imagens materiais que aparecem aos sentidos. Para Peirce,

<sup>44</sup>Uso som como exemplo porque pretendo mostrar a capacidade deste produzir imagens, mas é preciso ressaltar que também interessa aqui a capacidade da trilha visual do audiovisual imaginar música.

Um signo intenta representar, em parte, pelo menos, um objeto que é, portanto, num certo sentido, a causa ou determinante do signo, mesmo que o signo represente o objeto falsamente. Mas dizer que ele representa seu objeto implica que ele afete uma mente de tal modo que, de certa maneira, determina, naquela mente, algo que é mediatemente devido ao objeto. [...] (CP 6.347). (apud SANTAELLA, 2005, p. 43)

Desse modo, podemos compreender um som de piano saindo de alto-falantes como um signo que representa um som de piano no mundo, e o representa falsamente. Ainda assim, o que geralmente ocorre é que ignoramos esta questão e fruímos da reprodução técnica como se não fosse: “Nesse contexto, Peirce define *representar* como 'estar para, quer dizer, algo está numa relação tal como um outro que, para certos propósitos, ele é tratado por uma mente como se fosse aquele outro' (CP 2.273)”. (NÖTH; SANTAELLA, 2009, p.17)

Está posto aí uma primeira distinção entre *imagem* e *representação*. Imagens podem ser representadas por outras imagens, como a reprodução técnica de um som de piano pode também representar a imagem visual do músico a executando, mas também não precisam necessariamente representar algo, como uma imagem abstrata ou um som sintetizado eletronicamente. Uma imagem pode estar no mundo apenas para si, não precisando estar em relação a algo. Já uma representação necessariamente precisa significar outra coisa. E um ponto comum é que ambas não necessitam referir-se a um objeto “real”, a algo que está no mundo.

Paralelo a Peirce, Ferdinand de Saussure desenvolveu a semiologia. Para o autor, “O signo lingüístico une não uma coisa e uma palavra, mas um conceito e imagem acústica. Esta não é o som material, coisa puramente física, mas a impressão (*empreinte*) psíquica deste som, a representação que dêle nos dá o testemunho de nossos sentidos [...]” (1993, p. 80). O conceito é uma ideia geral do que o signo possa representar, como *árvore*, não uma árvore específica, mas qualquer árvore, e está atrelado à uma imagem mental de como deve ser pronunciado “árvore”. Embora Saussure estivesse pensando na língua, a noção de *imagem acústica* interessa aqui, pois compreende um evento sonoro convertido em memória, como imagem. Mais tarde, o autor mudou os termos. O conceito passou a ser o significado, e a imagem acústica, significante. Tal mudança foi feita para aumentar a amplitude de sua teoria. O modelo se encerra nessa bilateralidade: da união de ambos emerge o signo. Um movimento importante que a semiologia faz, portanto, é o de excluir o objeto de referência do modelo.

É uma interpretação diferente da que Santaella e Nöth introduzem, pois o signo deriva de duas imagens. Tanto os signos quanto as imagens que os formam não necessitam representar algo que está “no mundo” na semiologia.



O que haja de idéia ou de matéria fônica num signo importa menos que o que existe ao redor dêle nos outros signos. A prova disso é que o valor de um termo pode modificar-se sem que lhe toque quer no sentido quer nos sons, unicamente pelo fato de um termo vizinho ter sofrido uma modificação (ver p. 134 s.). (SAUSSURE, 1933, p. 139)

Essa é uma visão interessante para esta dissertação, pois como vimos no caso do vídeo *MC Jeremias* no primeiro capítulo, as remixagens evidenciam o caráter mágico das imagens técnicas, e nos fazem lembrar que elas não representam o mundo, e sim conceitos do mundo. Além disso, a produção de sentido nos vídeos musicais para *web* se dá também na rede de relações entre imagens que são rearranjadas pela mixagem. Saussure também atesta que tal rede se dá por oposição, diferenças e valores dos signos, e aqui podemos estabelecer até uma relação com a Música: uma melodia se dá através da diferença de sonoridade entre as notas, o que é considerado um intervalo musical. É pela diferença que a linguagem musical se constitui, e assim o é também no audiovisual. Saussure compreende as relações que se dão no tempo como relações sintagmáticas: é o eixo diacrônico do audiovisual e da Música. Mas há ainda um segundo eixo, que é o das relações paradigmáticas, ou associativas, como Saussure as denomina, que “une termos *in absentia* numa série mnemônica virtual. (1993, p. 143). Esse grupo de palavras é da ordem do possível, porque na linguagem escrita só é possível expressar-se linearmente. Mas no caso da Música, através da polifonia, e no audiovisual, através de elementos diversos que podem ser considerados isoladamente (trilha visual e sonora, cor, narrativa, movimento, ritmo, enquadramento, montagem), há, além de um eixo diacrônico, um eixo sincrônico em que se podem estabelecer sentidos no encontro entre imagens (visuais, sonoras e/ou musicais) diferentes e simultâneas.

Um caso interessante de se notar isso é Eisenstein, que dirigiu filmes entre 1923 e 1948, sob o regime do cinema clássico e do formalismo russo. Em sua obra, a preocupação residia em que a montagem suprisse a produção de uma imagem que sintetizasse a proposição do filme – que fosse capaz de produzir conceitos. Em uma sequência do filme *Outubro* (1928), o diretor alterna dois planos: no primeiro, o primeiro ministro Alexander Kerenski com toda sua pompa, sapatos lustrados e roupa impecável; no segundo, um pavão animatrônico ostenta suas plumas. Da soma dos dois emerge um conceito, uma crítica aos homens de poder.

*A representação A e a representação B devem ser selecionadas entre todos os aspectos possíveis do tema em desenvolvimento, devem ser procuradas de tal modo que sua justaposição – isto é, a justaposição desses próprios elementos e não de outros, alternativos – suscite na percepção e nos sentimentos do espectador a mais completa imagem do próprio tema. (EISENSTEIN, 2002, p. 18)*

Na soma de vários pares de planos justapostos, “de um modo ou de outro, a série de ideias é montada, na percepção e na consciência, como uma imagem total, que acumula os elementos isolados” (EISENSTEIN, 2002, p. 21). É nessa ordem de significação que Eisenstein estava interessado. Queria encontrar uma linguagem específica do cinema. No seu cinema, o significado das cenas emerge entre imagens (tanto sincrônica quanto diacronicamente). Esse significado, por sua vez, também pode ser considerado como uma imagem mental (se nos remetermos a Nöth e Santaella), que nasce da justaposição das representações A e B na percepção do espectador.

Essa ideia de *entre imagem* é importante nessa dissertação, pois a maioria das vezes em que se reconhece uma imagem-música nas análises (ver capítulo 4), ela deriva do choque entre duas outras imagens. Algumas vezes, isso pode se dar sincronicamente, entre uma imagem visual e outra sonora, às vezes diacronicamente, como uma melodia que se desenvolve no tempo, mas também, e principalmente, entre o audiovisual como um todo e a virtualidade da Música.

Para compreender melhor a natureza dessas imagens, é preciso evocar dois autores, os quais serão decisivos para a compreensão do que é imagem para esta dissertação: Flusser e Bergson. O primeiro nos ajuda a compreender o caráter mágico das imagens técnicas, aquelas produzidas por aparelhos. Mas o que são “imagens técnicas”? Flusser nos ajuda a decifrá-las. O central aqui é entender que “O que vemos ao contemplar as imagens técnicas não é 'o mundo', mas determinados conceitos relativos ao mundo [...] (2011a, p. 31)”. Para termos uma noção da distância que essas imagens estão da coisa representada em si, ou ao menos da representação primeira da coisa, basta lembrar que elas estão um nível acima de abstração em relação às imagens tradicionais. Para o autor, as tradicionais, como a pintura, primeiro transformam um acontecimento em cena (abstração primeira); em seguida, suprimem a terceira dimensão e criam conceitos de profundidade para poderem representar a do próprio mundo em uma superfície bidimensional (abstração segunda). A imagem técnica abstrai mais uma dimensão, transformando a imagem tradicional em texto, em conceito. No caso das imagens digitais, que são as que tratamos nesta dissertação, esse texto é unidimensional: o código binário (abstração terceira). E cada cálculo que o computador faz para desenvolver tais textos ocorre num plano nulodimensional:

As imagens da imaginação até hoje são bidimensionais porque foram abstraídas do mundo, digamos, quadridimensional; e as imagens da nova imaginação são bidimensionais porque foram projetadas por cálculos adimensionais (*nulldimensional*). O primeiro tipo de imagens faz a mediação entre o homem e seu

mundo; o segundo tipo, entre cálculos e sua possível aplicação no entorno. O primeiro significa o mundo; o segundo, cálculos (FLUSSER, 2007, p. 172-173)

Ao ser decodificada, ela converte os cálculos em uma representação de uma imagem tradicional, como as que vemos na tela da televisão ou em nossos computadores. Sinteticamente, “[...] as imagens tradicionais imaginam o mundo; as imagens técnicas imaginam textos que concebem imagens que imaginam o mundo. Essa condição das imagens técnicas é decisiva para o seu deciframento” (p. 30). Imagens técnicas, para Flusser, é a possibilidade de se produzir um pensamento em superfície, o que para ele seria o principal indício de que estaríamos vivendo a transição da História, de pensamento linear, para a Pós-História, de pensamento em superfícies. Outra diferença é que as imagens tradicionais têm como função predominante a de representar o mundo, enquanto que as técnicas a de elaborar conceitos.

Já Henri Bergson não chegou a conhecer as imagens técnicas digitais, mas sua teoria muito se presta para a compreensão de tais imagens. O filósofo compreende por imagem “uma certa existência que é mais do que aquilo que o idealista chama uma representação, porém menos do que aquilo que o realista chama uma coisa - uma existência situada a meio caminho entre a 'coisa' e a 'representação'” (1990, p. 1). É algo que está *entre*, e são as relações entre elas que atualizam incessantemente as coisas, restituindo-as de movimento. Para o filósofo, o universo todo é um conjunto de imagens. Nosso corpo é capaz de percebê-las e interagir com elas, devolvendo-as ao fluxo de um modo diferente de como elas foram percebidas. Para o autor, tudo está em movimento, ou melhor, não há coisas que mudam: há mudança. Bergson deixa de lado a discussão entre coisa e representação da coisa para pensar nas relações apenas entre imagens: “Toda imagem é interior a certas imagens e exterior a outras; mas do conjunto das imagens não é possível dizer que ele nos seja interior ou que nos seja exterior, já que a interioridade e a exterioridade não são mais que relações entre imagens” (1990, p. 16). Esse seu modo de pensar é muito produtivo para compreender como o audiovisual produz sentidos apenas através da articulação entre imagens.

A matéria “é uma imagem, mas uma imagem que existe em si”. Longe de ser estática, essa imagem que existe em si é um estado de coisas que não cessa de mudar, ela é matéria-escoamento sem ponto de ancoragem, nem centro de referência assinaláveis. Assim, de início, [o livro] *Matéria e Memória* traçava um plano em que se identificavam a matéria, a imagem e o movimento. Em outras palavras, em que a matéria-escoamento é imagem-movimento. (FORZANARI, 2010, p. 94)

As imagens produzidas pelas telas de vídeo (considerando aqui televisores e monitores de computador) se comportam de um modo muito similar ao entendimento de Bergson. Se

“[...] o cinema não trabalha verdadeiramente com o tempo enquanto categoria significativa, uma vez que a impressão de movimento que ele nos dá é um efeito de percepção (denominado efeito phi em psicologia)” (MACHADO, 1993, p. 101), o vídeo é completamente diferente. Ao invés da imagem ser projetada numa tela, a própria tela é constituída por centenas de milhares de pontos luminosos que são atualizados em altíssima velocidade. Segundo Machado,

[...] a imagem eletrônica não consiste em outra coisa que um conjunto de linhas sobrepostas e sucessivas, cada uma delas constituída por um número dado de pontos elementares de cor, que se juntam num quadro para formarem um painel de retículas. À maneira da técnica pictórica do mosaico. Ela é também uma imagem iridiscente, uma imagem-luz, em que a informação plástica coincide com a fonte luminosa que a torna visível. Tecnicamente, a imagem eletrônica se resume a um ponto luminoso que corre a tela, enquanto variam sua intensidade e seus valores cromáticos. (1993, p. 113-114)

Esse detalhe tem uma enorme significância no que diz respeito ao estudo do tempo no audiovisual. Enquanto que a “menor unidade” do cinema são imagens estáticas equidistantes em tempo e espaço (os fotogramas) que criam uma ilusão de movimento, no vídeo,

[...] em cada fração de tempo, não existe propriamente uma imagem na tela, mas um único pixel, um ponto elementar de informação de luz. A imagem completa, o quadro videográfico, não existe mais no espaço, mas na duração de uma varredura completa da tela, portanto no tempo. A imagem eletrônica não é mais, como eram todas as imagens anteriores, ocupação da topografia de um quadro, mas a síntese temporal de um conjunto de formas de mutação. (MACHADO, 1989 apud MACHADO, 1993, p. 114)

Por esse motivo, pode-se ter um controle maior sobre a imagem, sendo possível abrigar todo tipo de imagem e fazer com que elas dialoguem. A textura do vídeo compactado é produzida por passagens: todo tipo de imagem pode passar por ela, fotografia, filme, infografia, e com técnicas como o cromaqui é possível incrustá-las, invertê-las, converter uma noutra. Não há mais imagem íntegra no vídeo, pois mesmo uma aparente fotografia estática na tela é composta pelo mosaico de pontos. Materialmente, fisicamente, ela não está lá. “Não há mais imagens simples, elas todas estão imbricadas. Cadeias de imagens, escravas umas das outras, sobre as quais perdemos todo o poder.” (PEIXOTO, 1993, p. 245). De fato, não há mais imagens que não sejam compostas de várias outras. Estamos no domínio do entre-imagens:

O “entre-imagens” é o espaço de todas essas passagens. Ao mesmo tempo absolutamente visível e secretamente imerso nas obras, flutuando entre dois fotogramas ou entre duas telas, entre duas espessuras de matéria ou entre duas velocidades, ele opera na intermediação das imagens. O entre-imagens é o lugar onde a paisagem contemporânea efetivamente se constitui (p. 239-241).

A proposta de Bergson é compreender todo o universo funcionando dessa maneira (para o autor, até nosso corpo é imagem), e o caminho que pretendo adotar nesta dissertação é o mesmo: se o modo como se produz imagens na trilha visual do vídeo facilita o entendimento da teoria bergsoniana, é preciso lembrar que o modo como se produzem imagens na trilha sonora se dá do mesmo modo (o que abordarei a seguir); e o modo como percebemos e nos relacionamos com o mundo também.

*Imagem*, nesta dissertação, portanto, é algo que não precisa representar algo que está no mundo, tampouco chega a ser uma coisa. Como foi dito em 1.3, ela sempre está relacionada a um virtual; além disso, ela mesma também tem uma virtualidade que lhe é própria. O aparecimento de uma imagem movimentada as estruturas da virtualidade da qual essa imagem deriva. Essa imagem pode ser um atual, que, ao atualizar-se, realiza materialmente uma parte do virtual; ou pode ser um possível, resultante da realização, que, por criar novas formas, conseqüentemente aumenta as possibilidades de atualização do virtual. A imagem, portanto, tem duas faces: uma virtual e não representacional, outra atual, necessariamente representacional (mesmo que represente ela mesma).

É preciso atentar também para os lugares onde se procurará tal ordem de imagem nesta dissertação. Na matéria, será na parte física do audiovisual digital: os códigos binários que, ao serem interpretados pelo computador, convertem cálculos em trilha visual e sonora. Há imagem antes de tal conversão (nos discos rígidos, DVDs, vinis e fitas cassetes), imagens visuais e sonoras, mas que são potenciais: elas não são ainda a manifestação plena, mas potencial de tais imagens. Há ainda uma segunda ordem de imagem, que é a da memória: composições de luz e som que irradiam e vibram das telas e caixas acústicas passam por nossa percepção e são convertidas em imagem, desta vez, imagens atuais mentais, que vão modificar, aumentar a complexidade do virtual. Isso é possível porque

Meu corpo é, portanto, no conjunto do mundo material, uma imagem que atua como as outras imagens, recebendo e devolvendo movimento, com a única diferença, talvez, de que meu corpo parece escolher, em certa medida, a maneira de devolver o que recebe. Mas de que modo meu corpo em geral, meu sistema nervoso em particular, engendrariam toda a minha representação do universo ou parte dela? Pode-se dizer que meu corpo é matéria ou que ele é imagem, pouco importa a palavra. Se é matéria, ele faz parte do mundo material, e o mundo material, conseqüentemente, existe em torno dele e fora dele. Se é imagem, essa imagem só poderá oferecer o que se tiver posto nela, e já que ela é, por hipótese, a imagem de meu corpo apenas, seria absurdo querer extrair daí a imagem de todo o universo. *Meu corpo, objeto destinado a mover objetos, é portanto um centro de ação; ele não poderia fazer nascer uma representação.* (BERGSON, 1990, p. 11)

Tal movimento – das imagens passarem por nossa percepção e serem devolvidas diferentes ao fluxo – é o que nos torna capazes de interferir no mundo.

Vale notar a diferença de leitura entre esta última e a que Santaella e Nöth (2009) apresentam na introdução de *Imagem...*: as imagens materiais são essencialmente voltadas para a nossa visão, como a pintura e a fotografia; as imateriais são imagens mentais, e podem abrigar diferentes tipos de representação, ou de serem criadas pela imaginação. Mas para Bergson não há diferença entre matéria e imagem. A partir dessa compreensão, a seguir vamos tratar da materialidade da trilha sonora do audiovisual de modo a pensar como ela pode produzir imagens.

### 3.2 A IMAGEM SONORA

A ideia de que o cinema era mudo no seu princípio reside no fato de que não havia um método de sincronização de registros fonográficos com a trilha visual do cinema (o rolo de filme sendo projetado). No entanto, os diretores sempre procuraram acrescentar sons de alguma forma. Podia ser com eventos na própria sala de projeção: músicos executando peças ao vivo ou um narrador que descrevia o que estava ocorrendo na cena. Ou podia ser com métodos visuais que sugerem sons, como ritmos mais frenéticos que remetiam a ritmos sonoros; com intertítulos que sugeriam o que as personagens conversavam; com a visualização de músicos e dançarinos ou objetos feitos para produzir sons, como um apito de fábrica. Não era o cinema, a sala de cinema, mas o filme, a película é que era muda, para sermos mais exatos.

Até pode-se considerar o ruído do projetor como parte do cinema, mas isso é muito discutível. Como não havia uma antessala com isolamento acústico para o projetor como as salas têm hoje em dia, o barulho incomodava muito as primeiras audiências, e foi esse um dos motivos da inserção de músicos nas salas, “[...] evitando que a atenção do espectador fosse desviada pelo barulho metálico da grifa. Como nova atenção, a preocupação dos exibidores era agradar o espectador, sem que este tivesse motivo de distração que não o filme” (MANZANO, 2003, p. 27). Sem dúvida esse ruído faz parte da sala de cinema (quem senta no fundo, mesmo nos dias de hoje, pode ouvi-lo), mas faz parte do filme, do audiovisual que estamos assistindo? Se fosse assim, não deveríamos considerar todos os ruídos da sala (as tosses, os espirros, a pipoca mastigada) como parte do filme? É interessante como Flusser

aborda esta questão. Para o autor, se os receptores “viram a cabeça em direção ao projetor, não é para libertar-se da ilusão, mas o fazem zangados, se o aparelho funcionar mal, e as sombras saltarem em vez de se movimentarem suavemente. Estão zangados, se a ilusão for desmascarada” (2011b, p. 84-85). E diz, ainda, que “o seu comportamento prova que *querem ser enganados*. (p. 85)

Portanto, os esforços tanto dos cineastas quanto dos receptores foi o de isolar o máximo possível esses sons da compreensão do filme sendo projetado. Sujeito e objeto estão bem discernidos aqui. É preciso diferenciar o que é do mundo do filme e o que não é. A música fílmica ajudou especialmente neste sentido.

Sentiu-se essa necessidade para poupar o espectador do estranhamento envolvido em ver efígies de pessoas vivendo, atuando e até mesmo falando, que eram ao mesmo tempo mudas. O fato de que elas são vivas e não-vivas ao mesmo tempo é o que constitui seu caráter fantasmagórico, e a música foi introduzida não para supri-los com a vida que lhes faltavam [...] mas para exorcizar o medo ou ajudar o espectador a absorver o choque. (PRENDERGAST, 1992, p. 3, tradução do autor)<sup>45</sup>

Nos primeiros vinte anos do cinema, o acompanhamento sonoro era feito por músicos que improvisavam ao vivo durante a exibição. Na maior parte das vezes, os músicos usavam partituras-coringa, peças que representavam a sensação de amor, drama, conflito, raiva, e trocava-se de uma para outra abruptamente durante a exibição, de acordo com o que era exibido na tela. Muitas vezes eles nem assistiam previamente o filme para terem uma noção do que iam tocar. Foi por isso que progressivamente o quarteto de cordas regido por maestro foi trocado pelo pianista solo, capaz de fazer transições mais suaves (e custar menos) (MANZANO, 2003). Dessa fase não há muito o que dizer da relação sincrônica entre música e filme, pois cada exibição era única. Mas é importante notar que, enquanto os diretores se preocupavam muito com questões de narratividade, enquadramento e tudo o mais que podiam fazer com câmeras e películas, a música e os músicos acompanhavam como podiam as imagens que corriam pelo projetor.

Portanto, houve a presença da música desde o início, mas ela era sempre usada como mera acompanhante das imagens. Mesmo com o advento de tecnologias de sincronização de película com algum tipo de registro fonográfico no final da década de 1920, a música seguiu exercendo o papel de ambientadora, de pano de fundo, e outros sons como a fala e ruídos não conseguiriam se autonomizar da trilha visual. É o que Deleuze propõe, quando lembra que “o

---

<sup>45</sup>“The need was felt to spare the spectator the unpleasantness involved in seeing effigies of living, acting, and even speaking persons, who were at the same time silent. The fact that they are living and nonliving at the same time is what constitutes their ghostly character, and music was introduced not to supply them with the life they lacked [...] but to exorcise fear or help the spectator absorb the shock.”

falado, o sonoro são ouvidos, mas como *uma nova dimensão da imagem visual, um novo componente*. É sob esse estatuto, aliás, que são imagem” (2007, p. 269). Então, mesmo com a possibilidade de sincronização e inserção de ruídos e falas, o cinema seguiria utilizando os sons e a música em favor da trilha visual. Os filmes da época exageravam nos diálogos, o que resultou também num “achatamento” da imagem visual. Progressivamente, a quantidade de planos fechados foi aumentando para dar ênfase aos diálogos, e a noção de perspectiva visual, que era muito evidente nos filmes mudos foi perdendo importância. Para compensar isso entravam os sons, que restituíam essa dimensão. Um *closeup* de uma pessoa acompanhado de sua voz com uma reverberação longa já nos situa: ela está em um ambiente fechado, com grandes paredes e pé direito alto. Se o timbre for seco, pode estar em um quarto pequeno. Se um plano aberto em seguida frustrar essa expectativa, ou o diretor quer provocar a audiência, ou o técnico de som é muito ruim. De qualquer forma, devemos reparar que mesmo aqui o timbre sonoro é usado para imaginar como um ambiente deve aparecer visualmente (e mais uma vez nos colocamos no lugar do pensador dentro da caverna).

Estudos relacionados ao uso do som aparecem na obra de Eisenstein. À época (década de 1940), o som sincronizado no cinema já havia se popularizado, mas ainda se lê em sua obra “imagem” como algo exclusivo da trilha visual (da sequência de fotogramas) e “som” como exclusivo da trilha sonora (2002, p. 60). O conceito de montagem polifônica, que decorre do cinema sonoro, ainda é pensado para produzir na lógica do cinema clássico:

um plano é ligado ao outro não apenas através de uma indicação – de movimento, valores de iluminação, pausa na exposição do enredo, ou algo semelhante –, mas através de um *avanço simultâneo* de uma série múltipla de linhas, cada qual mantendo um curso de composição independente e cada qual contribuindo para o curso de composição total da sequência. (2002, p. 55)

Então, em Eisenstein, assim como no cinema clássico, os componentes sonoros passam a ser uma quarta dimensão da trilha visual. “No ponto em que estamos, o cinema continua a ser uma arte fundamentalmente visual” (DELEUZE, 2007, p. 286). Esse argumento de Deleuze difere bastante do senso comum do que é audiovisual, como apresentado inicialmente. Embora possamos dizer que o cinema clássico é audiovisual, a partir do autor não podemos negar que todos os seus elementos trabalham a favor do visual. Se há som, ele é submetido a preencher os espaços que faltam no visual, e por isso enriquece e muito o cinema clássico com as ideias de voz *off*, fora de quadro e de campo, trilha diegética e extradiegética. Chion contribui ainda dizendo que no cinema o som é vococêntrico: “É a voz que, na rodagem, é captada na tomada de som, que é quase sempre, de fato, uma tomada



de voz; e é a voz que se isola na mistura, como um instrumento solista, do qual os outros sons, músicas e ruídos, seriam apenas o acompanhamento” (2008, p. 13). No entanto, entre o final da década de 1950 e início de 1960, uma série de filmes contestariam essa fórmula: surge o cinema moderno. Costuma-se compreender o filme *Acossado* (1960) de Jean-Luc Godard como o pioneiro, pois o diretor optou por cortar trechos de planos, mas não no começo ou no fim, e sim no meio, causando uma supressão de tempo que produzia um salto, por isso nomeado de *jump cut* (também conhecido por *faux-raccord*). Mas o cinema moderno guarda também uma série de novidades para a trilha sonora. O próprio Godard procura novas possibilidades na montagem sonora, e provoca o espectador rompendo com hábitos do cinema clássico. A trilha sonora ganha autonomia em relação à trilha visual, não necessariamente tendo que trabalhar para ela: torna-se possível pensar conceitualmente em uma *imagem sonora*.

o falado e o conjunto do sonoro conquistaram autonomia: [...] deixaram de ser um componente da imagem visual, como no primeiro estágio, tornaram-se imagem integralmente. A imagem sonora nasceu em sua própria ruptura, de sua ruptura com a imagem visual. (DELEUZE, 2009, p. 297)

É desta forma que pretendo compreender o som nos vídeos musicais para *web*: em sua capacidade de significar de forma heautônoma, sem precisar necessariamente completar algo que falta à imagem visual. Penso também que as imagens sonoras possuem dupla face: por um lado, interferem e significam conceitualmente (n)os vídeos, e o fazem mais ou menos nos moldes das imagens sonoras que Deleuze reconheceu no cinema moderno; por outro, interferem e significam tecnicamente (n)os vídeos, e o modo como o fazem será esmiuçado a seguir.

Como, de acordo com o que acabamos de propor, a partir do cinema moderno se tornou possível pensar em uma imagem sonora conceitual, pretendo agora compreender o caráter técnico de tais imagens. Tratarei primeiro daquelas representadas tecnicamente, antes de serem decodificadas e soadas pelas caixas acústicas, e em seguida das imagens que nossa percepção produz a partir dos sons que ouvimos.

Até antes da era digital, sempre houve alguma diferença técnica entre a captação e/ou armazenamento do visual e do sonoro. No cinema, os fotogramas, como se sabe, comportam as imagens que vemos projetadas na tela. O que não é tão reconhecido é onde e como o som é registrado. Ao lado e ao longo dos fotogramas pode-se ver um fino rastro ondulatório (a figura 1 mostra um exemplo de um só canal, monaural): são ondas sonoras convertidas em imagem,

que ao serem lidas por um aparelho, são convertidas em pulso elétrico que faz vibrar as caixas acústicas.

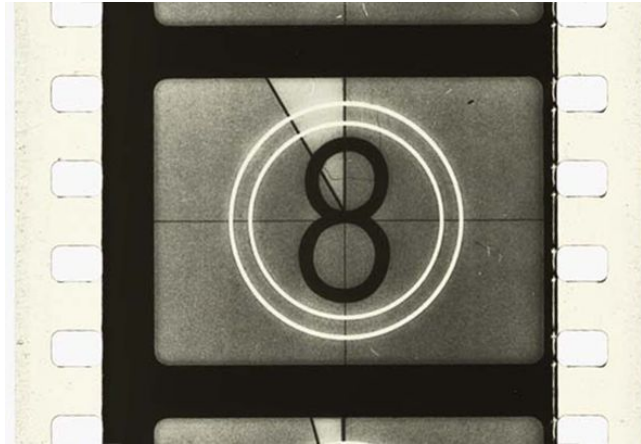
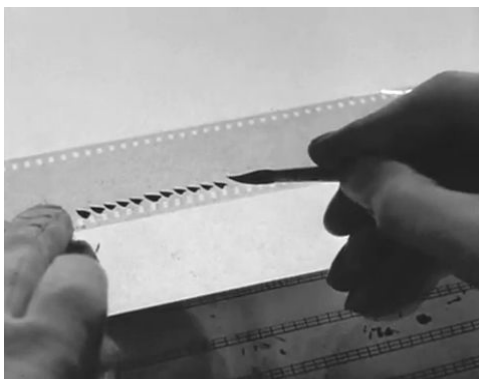


Figura 1<sup>46</sup>

Percebendo os padrões ondulatórios das imagens sonoras contidas nas películas, Norman McLaren ficou reconhecido por pintar diretamente na trilha sonora, de modo a produzir sons sem a necessidade de registrá-los previamente. Como exemplifica o documentário *Pen Point Percussion* (dirigido por Don Peters e Lorne Batchelor, 1951), ele sincronizava padrões vibratórios pintados na estreita faixa da trilha sonora (*Frame 5*) com desenhos que fazia na trilha visual (*Frame 6*), também pintando diretamente na película, o que resultava em filme produzido sem câmeras nem microfones.



*Frame 5*



*Frame 6*

Também pensando em visualizar os sons de algum modo, vários músicos modernos produziram partituras diferentes daquelas a que estamos habituados. Muitas delas exibiam desenhos no lugar da pauta e da notação. Alguns deles pediam para que músicos os

<sup>46</sup>Fonte: [http://www.exploremcallen.com/Images/events/film\\_reel.jpg](http://www.exploremcallen.com/Images/events/film_reel.jpg)

interpretassem, outros utilizavam softwares que convertiam as imagens em sons. Mas o que eles não conseguiam era fazer a união, a simultaneidade e a sincronicidade das imagens visuais e sonoras. Ou contemplava-se a partitura como uma arte figurativa, ou ouvia-se a música que resultava da interpretação destas imagens (a distinção schopenhaueriana discutida por Flusser). Nesse sentido, só McLaren conseguiu realizar tal objetivo, pois o material com que compunha o visual e o sonoro é o mesmo e pode-se assistir a ambos simultaneamente. E mais do que isso, separar as trilhas para contemplá-las individualmente destituiria todo o sentido de seu trabalho.

Ainda que McLaren faça imagens tradicionais tanto na trilha visual como na sonora, o fato de elas adquirirem movimento através do cinematógrafo as transformam em *imagens técnicas*. Seu ato de pintar diretamente na película implica na produção de imagens tradicionais; o ato de serem projetadas, adquirindo movimento visual e sonoro através de um aparelho, as transformam em técnicas<sup>47</sup>.

Mas só porque Flusser construiu sua teoria usando a caixa preta da máquina fotográfica como metáfora não significa que ela só sirva para explicar questões referentes a visualidades. Vejamos como é possível pensar em imagens técnicas do som.

Quando se desenvolvem métodos de registro digital de vídeo, pela primeira vez imagens visuais e sonoras são armazenadas usando um mesmo código (sem considerar a técnica de McLaren, que é a exceção da regra): o binário. E aqui é o ponto fundamental da discussão, porque serão de vídeos digitais que iremos tratar.

Mas aí reside a dificuldade que estamos enfrentando. Ora, se a linguagem binária codifica imagens linearmente em um arquivo de dados, e as representa após um *software* decodificar o arquivo, não teríamos nesse arquivo, como banco de dados, “imagens de imagens”? E no caso do som, “imagens de sons”? Parece que faz-se necessário recuperar e analisar estes movimentos de codificação. Quando buscamos por *sons* em audiovisuais, só haverá *som* num entre-imagens: as imagens que o computador cria do som capturado por um microfone via código binário, por exemplo, e as imagens que nossa mente produz dos sons que ouviu. Entre um e outro há o fluxo, o tempo, a vibração incessante do ar, num movimento que se dá *depois* da caixa de som vibrar e *antes* do tímpano de nossos ouvidos serem vibrados. Dessa vibração do ar, nossa percepção extrai imagens que as representam. O tempo todo, falamos de *imagens sonoras* como representações visuais (espectogramas) ou como representações mentais. Para a tradição kantiana isso parece óbvio: não podemos falar das

---

<sup>47</sup>O termo “técnica” em Flusser não está relacionado com o *techné* grego, da arte de saber fazer, mas com o paradigma da Pós-História.

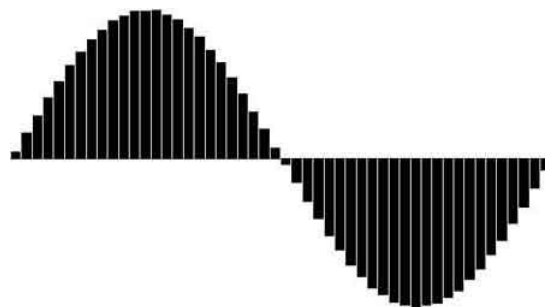
coisas, apenas das representações que temos delas. Mas porque então essa diferenciação confusa da trilha visual e sonora, chamadas de *imagem* e *som*? É preciso esclarecer essas questões não só no visual, onde essa questão é mais clara e estudada, mas também no sonoro. Antes de tudo, é preciso lembrar que

o som físico da acústica não é o som percebido pela sensibilidade. Esta distinção foi discutida com clareza por Duchez (1991: 49-50): “O *som físico* é um fenômeno energético, que consiste de oscilações regulares (quer dizer, de alternâncias de compressões e dilatações desse meio elástico); o *som percebido sensível* é um fato de consciência, possuindo com o som físico que o condiciona relações estreitas, mas que não são nem absolutas, nem constantes (pois elas dependem de propriedades fisiológicas da orelha e das modalidades psicológicas da audição).” (SANTAELLA, 2005, p. 108)

Vale lembrar que o som também é tridimensional, e passa pelas mesmas ordens de abstração: uma banda que grava uma canção no estúdio se desconstrói por completo: toca um instrumento por vez, utiliza outros equipamentos que não costuma usar ao vivo, adiciona reverberação de ambiente simulada, tudo para que a música gravada soe o mais natural possível. Convertem, portanto, um acontecimento em cena. Todos os sons captados são convertidos em pulsos elétricos, que ao passarem por um osciloscópio (*Frame 7*) se transformam em imagens bidimensionais, representando na vertical a intensidade sonora, e na horizontal a frequência; essa bidimensionalidade visual já não tem relação alguma com a “altura” e “largura” dos sons que ouvimos: não se trata de posição espacial, mas de tempo de duração (eixo horizontal) e volume (eixo vertical). A terceira abstração ocorre quando as curvas de onda sonora são convertidas em ondas serrilhadas: ocorre um micro fracionamento do eixo horizontal, criando colunas de alturas diversas (*Frame 8*). Cada coluna responde sozinha a uma intensidade sonora vertical (unidimensionalidade), e, quando varridas uma por uma, geram ondas similares às vistas no osciloscópio, e então são convertidas novamente em pulsos elétricos, para que enfim as caixas acústicas façam vibrar o ar e escutemos os sons.



Frame 7<sup>48</sup>



Frame 8<sup>49</sup>

<sup>48</sup>Extraído do filme *Pen Point Percussion*, (Don Peters e Lorne Batchelor, 1951).

<sup>49</sup>Fonte: <http://www.making-music.com/digital-audio.html>

Nesse caso descrevi como é uma onda sonora digital. Aquele gráfico do *Frame 8* “não existe”, ou melhor, não pode ser visualizado como as ondas visuais que o osciloscópio (*Frame 7*) apresenta. Ele é uma representação para que se compreenda como um arquivo em código binário converte pacote de dados em pulsos elétricos, e então convertidos em pulsos sonoros por caixas acústicas. É diferente do vinil, por exemplo, cuja trilha sonora é composta de sulcos que, ao serem lidos pela agulha, são convertidos em pulsos elétricos. Os sulcos na trilha são ondas que seriam análogas às ondas produzidas pela fonte sonora original. Nesse sentido, ambas são imagens técnicas, porque representam conceitualmente um som que “está no mundo”.

O objetivo neste capítulo era problematizar diferentes modos de pensar na trilha sonora como capaz de produzir imagens assim como trilha visual também o é, e estou fazendo isso pela perspectiva bergsoniana que esta dissertação adota. Uma vez que entendemos o universo dos vídeos musicais para *web* como um composto de imagens, podemos avançar para o desenvolvimento metodológico da dissertação.

## 4 PERCURSOS METODOLÓGICOS PARA O ESTUDO DA IMAGEM-MÚSICA NOS VÍDEOS PARA A *WEB*

### 4.1 O VIRTUAL DA MÚSICA NOS VÍDEOS PARA *WEB*

Este capítulo pretende pensar o que é Música, especificamente o que ela é para e neste projeto. Responder o que seja Música talvez seja impossível e completamente descabido, ainda mais depois de todo o desenvolvimento que se deu para a expansão de suas fronteiras nos últimos cem anos, de Schoenberg à música eletroacústica. Não se pode mais responder dizendo que são “sons organizados em ritmo, melodia e harmonia”, porque isso só responde à música tonal de uma época. Lúcia Santaella teve de enfrentar essa dificuldade quando redigia seu livro *Matrizes da linguagem e pensamento* (2005, p. 20), no qual defende a hipótese de que os três tipos de linguagem, verbal, visual e sonora, seriam as três grandes matrizes lógicas da linguagem e pensamento. No caso, para a linguagem sonora, o primeiro movimento da autora foi o de tentar relacionar o ritmo com o ícone peirceano, a melodia com o índice e a harmonia com o símbolo. No entanto, ao se deparar com as composições modernas, Santaella percebeu que

[...] uma classificação da linguagem musical nos componentes do ritmo, melodia e harmonia tem um caráter tão bem recortado quanto o das músicas a que ela é capaz de ser aplicada, isto é, o universo dos sistemas musicais que se utilizam de convenções para guiar a composição. Isso significa que ficaria excluída dessa classificação mais limitada a grande maioria das desafiadoras composições da música concreta, eletrônica e eletroacústica [...]. (2005, p. 87)

O funk carioca, por exemplo, é um estilo que na maior parte do tempo ignora a harmonia. Em algumas peças de Stockhausen, como *Carré* (1959-60), a transição entre algumas notas é tão longa que perdemos a noção de ritmo (cf. SCHAFER, 2001, p. 164). Depois de todas as vanguardas que surgiram, fica cada vez mais difícil responder tal questão. Algumas pessoas tentaram responder de outras maneiras, claro que não definindo de uma vez por todas o que seja Música, mas procurando alguns elementos de fundo. Para Schafer (2001), qualquer evento sonoro pode ser música, inclusive a paisagem sonora da cidade onde moramos. Para Wisnik (1999), é qualquer coisa que ocorra no tempo e que tenha pulso, sendo este também a unidade mínima de uma onda sonora. Um pulso frequente e muito rápido pode se transformar em um som com altura melódica, e aí, deste ponto de vista, ritmo, harmonia e

melodia, enquanto eventos sonoros derivam todos do pulso. Mas no final das contas, dizer o que seja Música é circunscrever as possibilidades do que pode vir a sê-la. Talvez por isso o dicionário Grove de música (SADIE, 1994) não tenha um verbete destinado a esse termo em suas mais de mil páginas, preferindo definir termos como “música das esferas”, “música de câmara”, “música folclórica” e assim por diante.

Não pretendo responder tal pergunta, mas sim pensar as condições de possibilidade da Música atualizar-se em vídeos para *web*: que elementos desta virtualidade se atualizam com mais recorrência. Se voltarmos à pergunta *o que é Música?*, o primeiro movimento a fazer é pensar o termo “é” como indicador de movimento, de criação incessante de novas formas a partir da matéria Musical. Nos termos de Bergson, poderíamos pensá-la como uma bola de neve, que incha na medida em que avança, acumulando-se de memória (2006c). A Música não para de agregar a si cada vez mais informação. Sob essa perspectiva, podemos entendê-la como uma virtualidade.

Se tentássemos definir o que é a virtualidade “Música”, ela certamente resvalaria dos nossos termos. Acontece que ela (como todas as coisas) está sempre diferindo de si; as definições mais aceitas há cem anos não são as mesmas de hoje. É como se houvesse uma coleção gigantesca de imagens que remetem à “Música”, e que estão constantemente se metamorfoseando e aumentando em quantidade. Cada definição, cada imagem dessas é uma atualização de sua virtualidade, uma figura, uma representação. Em cada atualização há um pouco de Música nela impressa, e a cada novo atual sua virtualidade é reconfigurada. Os motivos que levam a esse movimento incessante são vários, mas mesmo quando ouvimos a uma música novamente, não será a mesma da primeira vez, pois estaremos com um diferente estado de espírito.

A ideia de virtual em Bergson também é contemplada por Deleuze, que aproxima o conceito ao de estrutura. Talvez possamos pensar a Música como uma estrutura que se atualiza nos seus elementos. Numa estrutura, coexistem “Todos os elementos, as relações e valores de relações, todas as singularidades próprias ao domínio considerado” (2005, p. 250). Estaríamos tratando da Música em sua *totalidade*. E como não somos capazes de perceber o todo dessa estrutura, ela é dotada de uma invisibilidade. Quando se tornam atuais, isto é, quando algo ou parte de sua totalidade se imprime em determinada materialidade, torna-se visível, já não é mais apenas a estrutura. Mas por ter algo dela, ela torna-se presente pela ausência. As estruturas seriam, portanto, da ordem do virtual.

Uma canção só é capaz de nos comunicar algo porque primeiro decodificamos as vibrações sonoras a partir da memória do que consideramos Musical – traduzindo-as em

imagens musicais, portanto. Essas imagens não são sonoras, tampouco visuais. São memórias: as referências e representações que temos do que entendemos por Música, que se somam ao que estamos percebendo para ali tentar encontrar certos elementos estruturais na canção. É só assim que podemos reconhecer um ritmo musical, porque nossa memória, dotada de imagens que tomou do primeiro compasso, reconhece nos seguintes a repetição dos elementos do primeiro.

Como apresentado no capítulo 3.1, estou considerando imagem de acordo com o conceito de Bergson, como algo que está entre a coisa em si e a sua representação. Ao entender a Música como virtualidade, estaríamos considerando todas as imagens dela que existem, e que estão em constante atualização. A Música, como sugiro nessa pesquisa, é capaz de se atualizar em qualquer imagem, e o faz especialmente em nosso corpo porque ele funciona como centro de ação: “Meu corpo é portanto, no conjunto do mundo material, uma imagem que atua como as outras imagens, recebendo e devolvendo movimento, com a única diferença, talvez, de que meu corpo parece escolher, em certa medida, a maneira de devolver o que recebe” (BERGSON, 1990, p. 11). Assim, ao memorizar os princípios básicos da composição musical (atualizações da Música), seríamos capazes de produzir músicas. Não são criações *do zero*, mas uma restituição de movimento de um determinado conjunto de imagens musicais que meu corpo selecionou, agrupou e atualizou em forma de canção.

As imagens que o audiovisual utiliza não são como as de nossos corpos; são imagens técnicas. Não trataremos aqui, portanto, de representação musical, nem de uma ontologia do audiovisual ou da Música, mas sim como a Música se atualiza *através* das imagens técnicas do audiovisual, produzindo imagem-música.

E da virtualidade da Música, algumas imagens se manifestarão com maior força. Neste momento, pretendo circunscrever alguns elementos estruturais da Música com maior potência de atualização pelos vídeos que constituem o corpus. Vale notar que não estou fazendo aqui previsão ou procurando linhas gerais que poderiam prever tendências – isso seria anti-bergsoniano. Em *O Pensamento e o Movente* (2006b), o autor comenta que só é possível encontrar tais potencialidades *depois* que elas se realizaram. Que só a partir dos atuais podemos descrever e encontrar suas potencialidades em atuais anteriores, ou mesmo no virtual de tal coisa. “Digamos, portanto, que na duração, considerada como uma evolução criadora há criação perpétua de possibilidade e não apenas de realidade” (BERGSON, 2006, p. 15). E embora venha se falando na maior parte do tempo do par virtual/real, é preciso considerar que quando o virtual se atualiza (que, segundo Bergson, é quando o real acontece), este atual automaticamente configura novos possíveis para o virtual, o que serve aqui para



complementar a explicação dos gráficos apresentados no capítulo 1.3. Bergson, por sua vez, problematiza o modo como normalmente se pensa o par possível/real:

Se deixamos de lado os sistemas fechados, submetidos a leis puramente matemáticas, que são isoláveis pelo fato de a duração não os atingir, se consideramos o conjunto da realidade concreta ou muito simplesmente o mundo da vida e, com mais forte razão, o da consciência, descobrimos que há mais, e não menos, na possibilidade de cada um dos estados sucessivos do que em sua realidade. Pois o possível é apenas o real com, em acréscimo, um ato do espírito que repele sua imagem para o passado assim que ele se produziu. Mas é isso que nossos hábitos intelectuais nos impedem de perceber. (2006, p. 114)

Seguindo esta lógica proposta por Bergson, as potencialidades abaixo descritas são resultado da análise do compósito de vídeos que constituem o *corpus* (atualizações de imagem-música) – e apenas deles. Tais elementos colaborarão na compreensão das análises no próximo capítulo

Um das coisas que reparei durante a pesquisa exploratória é que raramente ocorre uma prática singular de apenas um vídeo. Assim, a ideia de autor se enfraquece aqui. Na *web* quase sempre é possível reconhecer uma técnica sendo replicada em vários vídeos de usuários diferentes. Além disso, muitas das técnicas já apareciam em potência em vídeos anteriores à *web*, como é o caso de *Amateur* de Lasse Gjersten, cuja técnica já havia sido abordada pela dupla Godley & Creme na década de 1980. Talvez isso possa ser uma comprovação flusseriana de que a imagem-música é do aparelho, e não de seus usuários. A noção de direito autoral perde seu valor, tanto nos vídeos que os usuários utilizam e remixam, sem permissão, imagens de arquivo com direitos reservados, quanto nos vídeos cujos autores captam as próprias imagens. Eles sabem que possivelmente seu conteúdo será reutilizado por outros usuários. Então, praticamente todo vídeo é resultado da horizontalização da produção de imagens técnicas.

Na maioria dos vídeos, a Música manifesta três dos seus alicerces<sup>50</sup> mais potentes: o ritmo, a melodia e a harmonia, e isso é assim porque estão inseridos na cultura ocidental e tonal, o que permitirá estabelecer paralelos entre linguagem e Música, como Santaella concluiu: “[...] o paralelo entre a língua e a música só se presta à música tonal ou quaisquer outros tipos de músicas convencionais. Tudo que veio depois do tonalismo não tem mais condições de ser comparado ao verbal” (2005, p. 101-2). Só há dois tipos de vídeo em que ela não se manifesta plenamente: nos *Shreds*, onde ela é sugerida, pois por mais mal que uma pessoa toque um instrumento como uma guitarra, as notas sempre serão da ordem do sistema

---

<sup>50</sup>“Alicerce” pode servir como analogia àquilo que funda e sedimenta as estruturas.

tonal (ou atonal, serial, mas de qualquer forma inserida na lógica do tonalismo), e nos vídeos que imaginam uma execução audiovisual de *4'33"* de John Cage<sup>51</sup>.

A estrutura da canção popular também aparece forte nestes vídeos. As imagens-música que resultam da montagem dos vídeos na maior parte das vezes estão inseridas na cultura da música popular. Há a recorrência de gêneros musicais contemporâneos de sucesso ao redor do mundo: eletrônica, rock, pop, funk, rap e *hip hop*. Isso pode ter relação com o fato de que os estilos que encontrei utilizarem mais a tecnologia e práticas de *sampling*, de digitalização e outros efeitos eletrônicos. Há, portanto, a recorrência de efeitos sonoros desses estilos: o *auto-tunning*, o fracionamento do envelope sonoro<sup>52</sup>, inversões, *pitch shifting*, *scratching*, entre outros. Então o tonalismo e suas noções de tensão e relaxamento estão sempre muito presentes. Por tensão e relaxamento estou me referindo a vários elementos, aos quais destaco:

1) A relação entre nota fundamental (aquela que dá o tom da música) e sua dominante (a nota mais distante do tom). Para entender ligeiramente a noção de tônica e dominante, trago um exemplo. Vamos supor que quiséssemos compor uma canção no tom dó maior. Se pensarmos na escala musical partindo desse tom, temos Dó-Ré-Mi-Fá-Sol-Lá-Si. Se lembrarmos do piano, antes do Dó desta escala há um Si mais grave, e depois do Si também desta escala há um Dó mais agudo, então a nota mais distante do Dó não é Si, mas o Fá e o Sol, por estarem ambas a três tons de distância de um Dó. São elas, e especialmente o Sol, que ao serem tocadas provocarão uma sensação de tensão, de distanciamento, e que em seguida deverão ser resolvidas pelo retorno ao tom, no caso a nota Dó. Esse processo fundamentou toda a história do tonalismo, que encontrou seu ápice, em Bach e Wagner (cf. WISNIK, 1999). Todas as outras seis notas exercem funções diferentes em relação ao Dó. O Si exerce a função de sensível. Pela sua proximidade com a nota fundamental, ao ser executada exerce um “forte impulso em sua direção”, de acordo com o Dicionário Grove de Música (1994, p. 853). A nota Ré também trabalha neste sentido, pela sua proximidade. Mi e Lá tem uma função relativa, e não provocam tanta tensão como as outras notas, tendem a acompanhar a sensação de relaxamento promovido pelo Dó;

2) O desenvolvimento de relaxamento à tensão através da estrutura da canção popular, que em geral se dá mais ou menos desta forma: introdução + (verso + [ponte] + refrão).2 + [solo] + (verso + ponte + refrão) + final. A ponte e o solo estão entre colchetes por serem menos recorrentes. Ainda falando em geral, todos os trechos procuram se manter mais ou

<sup>51</sup>Para esclarecer porque uso o termo “imaginar” e não “representar” ou “executar” a peça *4'33"*, ver 4.5.

<sup>52</sup>O envelope sonoro compreende todos os estágios de um som propagado: o ataque, o corpo e o declínio.

Fracionar esse envelope digitalmente é descaracterizar o que ele tem de natural, como cortar o ataque de uma nota tocada em um piano e não ouvir o som do martelo atacando a corda.

menos próximos ao tom da música, exceto a ponte. Então o momento de maior tensão não é o refrão, como se poderia supor. Mas o refrão é o momento de liberação de energia, que será recarregada pelo retorno ao verso. Essas noções nem sempre estão relacionadas à relação tônica/dominante previamente mencionada. Muitas vezes se dão apenas pelo aumento da densidade sonora, como é o caso de *Smells Like Teen Spirit* da banda Nirvana (1991), em que quase toda a canção se dá em cima de um único *riff* de guitarra.

Penso ser importante também considerar como parte dessa virtualidade outras ordens de imagens que não são apenas relacionadas à linguagem musical. Quando iniciei esta pesquisa, estava muito mais voltado para compreender a manifestação da linguagem musical no audiovisual, mas não posso deixar de constatar da importância de elementos visuais que também compõem a virtualidade da Música e que aparecem com certa frequência nos vídeos aqui estudados.

Para listar alguns: a performance, a postura do *rock star*, do *rapper* ou do *MC*, a presença visual de instrumentos musicais, dança, efeitos típicos do videoclipe como falhas de decodificação e sujeiras. Há ainda a inserção de sons que simulam uma performance ao vivo: aplausos, gritos de plateia, reverberações.

Todos estes elementos são imagens visuais (e podem ser materiais ou mentais, se pensássemos a partir de Nöth e Santaella, 2009), e elas são utilizadas para musicalizar ao máximo possível o vídeo. Todas essas imagens remetem direta ou indiretamente àquilo que este projeto pretende compreender por Música.

Então o que se percebe é que a diferença que estes vídeos provocam na virtualidade da Música não se dá exatamente nas canções que emergem da trilha sonora, e sim no processo de composição. Interessa aqui então compreender os procedimentos de produção de novas formas através de matérias audiovisuais, bem como as potencialidades que deste processo emergem como imagem-música.

É esse recorte que pretendemos fazer para compreender a ação da Música nos vídeos para *web* que constituem o corpus: encarando-a como virtualidade, vamos encontrar e analisar as imagens que ela faz produzir através do audiovisual. Desse encontro, emerge a imagem-música, que é um atual, mas que ao mesmo tempo expande as possibilidades da Música. “É preciso aceitá-lo: é o real que se faz possível e não o possível que se torna real” (BERGSON, 2006, p. 119).

## 4.2 A MÚSICA COMO SISTEMA

Neste capítulo serão construídos os dois pontos de observação a serem adotados na análise, baseados na teoria dos sistemas de Nicklas Luhmann (2009). O autor propõe que se substitua a distinção sujeito/objeto pela de sistema/meio, o que modifica o modo como observamos o objeto de estudo. O primeiro ponto de observação trata especificamente de reconhecer como a Música (sendo sistema), diferencia-se de si no audiovisual (este sendo meio). Tal ponto de vista permitirá reconhecer imagens-música em vídeos para *web*. O segundo ponto de observação inverte a lógica do primeiro: compreender-se-á o audiovisual como sistema e a Música como meio. Desse modo, será possível analisar as imagens-música previamente reconhecidas, compreendo como elas complexificam o sistema Musical, agregando-lhe novas informações. Mas antes de esclarecê-los, introduzo uma observação empírica que me permitiu esquematizar esse procedimento de análise.

O blog Vinikento<sup>53</sup> promete que, ao se executar em simultâneo, em diferentes abas do navegador, o vídeo *Army of the pharaohs - Battle cry*<sup>54</sup> e *Fatties at the grocery store rant*<sup>55</sup>, se ouvirá um rap. O primeiro, com uma imagem estática do LP do grupo de rap Army of The Pharaohs, conta com uma música na trilha sonora intitulada *Battle Cry*, em sua versão instrumental, típico *sampler* sonoro utilizado de fundo para um MC cantar por cima. O segundo é apenas um rapaz segurando a câmera com a mão e enquadrando a si, reclamando de um problema que teve em um mercado, sem medir as palavras. Um não tem relação alguma com o outro, vale dizer. A sincronização dos vídeos, sugerida pelo autor do blog, provavelmente lhe veio à cabeça do mesmo modo que ocorreu com o autor de *MC Jeremias*. O jeito de falar do usuário NormsBarStool, que protagoniza o vídeo *Fatties...* se assemelha iconicamente com o de um *rapper*. A música só surge da intervenção do espectador, que sincroniza vídeos pela emergência de uma música potencial. Uma espécie de *mashup* audiovisual.

Quando assisti esta sincronização pela segunda vez, senti a necessidade de esperar alguns segundos para rodar o vídeo *Fatties...*, para que pudesse ouvir os primeiros segundos de *Battle Cry* fazendo as vezes de introdução, o que me deu a sensação de estar compondo, criando algo novo. Mas, considerando o que este capítulo irá propor, entendo que essa ideia

<sup>53</sup>Disponível em: <<http://vinikento.tumblr.com/post/767947227/how-to-make-a-rap-song>>. Acesso em: 15 dez. 2011.

<sup>54</sup>Disponível em: <[http://www.youtube.com/watch?v=QLx60TZ9\\_T8](http://www.youtube.com/watch?v=QLx60TZ9_T8)>. Acesso em: 15 dez. 2011.

<sup>55</sup>Disponível em: <[http://www.youtube.com/watch?v=QLx60TZ9\\_T8](http://www.youtube.com/watch?v=QLx60TZ9_T8)>. Acesso em: 15 dez. 2011.

de introdução não pode ter sido por mim inventada, como que por um *sujeito* que dá forma a um *objeto*. Ela surge de uma aparente necessidade de apresentar o tema, que transcende este rap, transcende todos os raps, faz parte no mínimo de toda a música popular: a Música é que pediu que eu compusesse uma introdução. E não falo da música que esta mixagem faz emergir, mas a Música, dessa vez não só como virtual, mas também como sistema, como na Teoria dos Sistemas de Nicklas Luhmann (2009).

Embora o autor pense em sistemas sociais, a teoria se estende para outros âmbitos, e assim podemos entender também a Música como um sistema. Basta, a partir da observação do mundo, distinguir o que *é* Música (Sistema) daquilo que *não é* (meio). Encontramos uma Música que contém estruturas, normas, operações, que está em movimento, modificado-se incessantemente, evoluindo e adaptando-se ao meio.

Esse meio pode ser o Mundo, mas essa é uma opção de observação vaga demais. Se limitarmos a observação ao audiovisual, fica mais simples fazer uma distinção.

O meio audiovisual pode provocar irritações na Música. Algumas dessas irritações podem ser trabalhadas internamente pelo sistema e então ser por ele apropriadas, o que aumenta a sua complexidade. Mas essa apropriação não é passiva. Os sistemas sempre se apropriam das perturbações nos seus próprios termos, se organizam por conta própria, num processo chamado autopoiesis. O sistema não evolui em conjunto com o meio; pelo contrário, o meio evolui por conta própria enquanto o sistema se adapta como for mais adequado para si. Por isso, ele “é, no entanto, aberto a esse mesmo meio, não em termos essenciais mas acidentais” (ILHARCO, 2003, p. 153).

Se pensarmos agora no processo composicional como meio, deixando de lado o audiovisual por um momento, isso quer dizer que não teríamos livre acesso a elementos sonoros para produzir música. Os sons estão por aí, livres, mas musicalizá-los, adequá-los ao sistema Musical implica em fazê-lo de acordo com o que a Música como sistema permite. É nesse sentido que um sistema é fechado: só pode se auto-produzir a partir de suas próprias estruturas. E o espectro de sons e sequências de sons já apropriados pela Música é enorme, tanto que inclui sons desafinados, fora do ritmo, permite até que se fale ao invés de se cantar, como é o caso do *Fatties at the grocery store rant*. Sistemas de alta complexidade estrutural como a Música “combinam em seu interior seleções estruturais que só podem obter por si mesmos, para que disponha de um repertório maior para a ação” (LUHMANN, 2009, p. 116).

Esses exemplos em até algum momento da história não eram considerados musicais. Mas eles irritaram tanto a Música que acabaram por ela sendo apropriados. Fica muito difícil, portanto, incluir nela um elemento extra-musical, pois que quase tudo (senão tudo) que é

realizado sonoramente vem sendo considerado musical, o que aumentou a complexidade da Música propriamente dita.

Na maior parte do tempo que pensamos ser criativos, quem realmente está pensando é a Música. Isso ocorre toda a vez que nos sentimos obrigados a fazer uma introdução para determinada música, de que ela tenha um refrão ou coisa do tipo. É porque não temos tanta liberdade quanto pensamos, e sim somente dentro do sistema fechado da Música. Nela podemos agir livremente escolhendo traços reconhecidamente musicais para produzir uma música nova. Mas dificilmente haverá uma renovação da Música. Não somos nós, não é o mundo que escolhe quais sons podem virar Música, ela própria o faz para se manter, expandindo-se: “seleciona acontecimentos (reduz a complexidade externa) para aumentar sua própria complexidade [...]” (MARCONDES FILHO, 2004, p. 428).

A relação de seleção entre o sistema Musical e o meio propõe um jogo. De um lado, a Música apresenta seus recursos até então disponíveis: ritornelo, semifusa, escala pentatônica, solo de pandeiro, grito, o que for que ela até hoje permite ser possível ser reconhecido como musical. Do outro lado estão os compositores: pode ser Chopin ou o internauta que sincroniza ou remixa vídeos. Uns estudam o jogo da Música e a acompanham. Outros a enfrentam, caso de John Cage e sua peça *4'33"*; caso também do Funk Carioca, que ignora as regras de afinação e o sistema tonal na canção popular; os exemplos seriam intermináveis, mas todos periféricos. Quando eles são aceitos, quando por fim a Música por irritação apropria-se deles, a Música se expande, mas por uma seleção feita por ela mesma, através de um *acoplamento estrutural* ao meio, que

seriam as pequenas frestas que o sistema abre para observar o ambiente externo hostil [...]. Esses acontecimentos são moderadamente apropriados como “irritações” no interior do sistema e geram fórmulas para descarregá-las, torná-las inócuas, aumentando a complexidade do próprio sistema. Isto o torna, opostamente, mais frágil diante das investidas externas. (MARCONDES FILHO, 2004, p. 430)

Os vídeos da *web* se configuram como espaços privilegiados para esse acoplamento. A situação se opõe ao pensamento então vigente na televisão e no cinema, que já venho comentando ao longo do trabalho, mas que não custa ressaltar, nos quais a Música aparece como ambientadora, como pano de fundo, ou como nos videoclipes, em que tendencialmente o visual se produz a partir de uma música preconcebida, idealisticamente *trilha sonora*.

No recorte dos vídeos para *web* que queremos evidenciar, a Música emerge de qualquer fluxo. Ela pode até mesmo se manifestar visualmente, porque também a composição

visual é pensada pelas vias das estruturas musicais em suas durações, movimentos e intensidades (hipótese defendida nessa dissertação).

Isso não quer dizer que essa musicalidade não existia antes em outros audiovisuais. Mas foram esses vídeos que evidenciaram o processo. Fica a pergunta: como devemos observar tal manifestação, se a proposta difere do que até então se vinha fazendo? Um primeiro ponto importante a esclarecer aqui: quando falava sobre a Música como um sistema para o jogo composicional<sup>56</sup>, o ponto de observação adotado é o de compreender a Música como sistema e o processo composicional como meio. Mas agora que estamos partindo para a problemática de como a Música interfere, *irrita* a linguagem audiovisual, forçando-a a produzir imagens-música, é preciso partir para outro ponto de observação: agora, a Música é meio e o audiovisual sistema. Talvez se faça necessário entender melhor a própria observação.

Quando falamos em observar, defrontamo-nos com uma primeira diferenciação: *observar/observador*. Observar é a operação, enquanto observador é um sistema que utiliza as operações da observação de maneira recursiva, como sequências para obter uma diferença em relação ao meio. (LUHMANN, 2009, p. 154)

Em Luhmann, a operação de observar se dá em dois momentos: distinguir e designar. A distinção mais comum na análise do audiovisual é a de separar para um lado o áudio (trilha sonora) e para outro o vídeo (trilha visual). Essa leitura normalmente compreende a trilha sonora como música, e, por conseguinte, compreende a trilha visual como um campo onde a Música não se manifesta (o visual no máximo mimetiza a música através de elementos sinestéticos ou correlatos, como o ritmo). De fato, historicamente o cinema e a tevê sempre usaram músicas como panos de fundo, mas não é assim que a Música se evidencia nos vídeos da *web* que compõem o corpus desta pesquisa. Neles, ela se apresenta em todos os fluxos e elementos audiovisuais. Essa não é uma manifestação *sonora*, mas *potencial*. Não reconhecemos essa ordem de manifestação por hábito, isolando-a da trilha sonora, ou entendendo a musicalidade das imagens visuais se influenciando pela Música muito mais do que *sofrendo* influência.

Luhmann designa o observador também como um sistema, capaz de observar a si próprio, o que ele denomina auto-observação, ou a outros sistemas, hetero-observação. Quando observamos a Música ou o audiovisual, não o fazemos a partir tão somente de nossa consciência, mas através de um sistema que faz hetero-observação. Fazemos parte deste sistema. A consciência observa por si, é claro, mas sempre como parte de um sistema que é

---

<sup>56</sup>Saliento que, quando estou me referindo ao processo composicional, não estou tratando exclusivamente da ideia de manifestações sonoras, mas também do processo de edição dos vídeos musicais para *web* aqui estudados.

quem propõe o que deve ser observado. O sistema, que deveria ser objeto de estudo do sujeito, vai-se autonomizando. Esse ponto de vista acompanha a metafísica, na dissolução da diferença entre sujeito e objeto, “mas ela mesma deve ser desconstruída na medida em que se vê que também os objetos podem ser sujeitos, isto é, observações realmente operantes, observações que observam o observador” (MARCONDES FILHO, 2004, p. 438). Não só o homem pode ser aquele que observa, mas agora qualquer sistema é apto a observar tanto a outros sistemas (inclusive a nós, humanos) quanto a si próprios.

É preciso considerar também o fato de que a Música se auto-observa durante seus processos. “[...] é o próprio sistema se cuidando, aparando as diferenças, reproduzindo-se, mantendo-se vivo. Não há nenhuma interferência humana aí. A observação é um modo específico de operação que se apoia na diferenciação” (idem, p. 431-2). Por isso que a Música é que se faz no homem, porque ela também é capaz de fazer hetero-observação e causar irritabilidade ao compositor, pois para este, que também é um sistema, a Música é meio. Várias leituras do que seja sistema ou meio podem ser feitas, como se pode ver, dependendo do ponto de vista do observador. Não é possível assimilar tudo, apenas escolher um lugar privilegiado para observar. É o que Luhmann chama de ponto cego, “um ponto de invisibilidade, que garante a unidade da diferença, não importando qual seja a distinção, uma vez que a unidade da diferença não é observável” (2009, p. 160). O ponto de vista que escolhermos sempre irá gerar um ponto cego, tanto do olho que não pode ver a si como de alguns pontos do observado que se tornam invisíveis.

Vamos tratar agora sobre os Sistemas: como eles funcionam, e como a Música pode ser entendida como tal. Uma das características fundamentais do sistema é sua instabilidade. Ele não pode ser estático, senão não poderia ser considerado autopoietico, por isso cada operação que ele realiza o modifica. “Essas operações precisam adequar-se ao movimento anterior (memória) e dirigem-se a um sentido interno do sistema, à sua expectativa. A operação aferra-se ao passado, projeta-se para o futuro servindo-se dessa memória, que é seletiva” (MARCONDES FILHO, 2004, p. 431). Essas operações surgem de uma instabilidade no sistema, provocado, por exemplo, por uma irritação que o meio produziu, e tratam de tentar restabelecer a ordem no sistema. O resultado dessa operação pode ser o da apropriação ou expulsão do elemento que irritou o sistema. “As irritações remetem às estruturas e só são aceitas se tiverem que ver com as expectativas do sistema: é a partir destas que configura-se uma perturbação e faz com que a autopoiese reaja, identificando-a ou recusando-a [...]” (MARCONDES FILHO, 2004, p. 428). São as estruturas que dão condição para que a *autopoiesis* possa ocorrer:



[...] o mínimo que se pode pensar é que uma estrutura não se constrói à imagem de uma *coisa* composta de elementos que possam ser reunidos. A especificidade das estruturas reside, antes, no fato de que elas constituem um processo de repetição, no sentido de que uma estrutura simula situações que entende como repetição. (LUHMANN, 2009, p. 117)

Nesse cenário, voltamos à história da necessidade que canções populares têm de reapresentar o tema, de ter refrões. Estas são necessidades que se repetem, portanto, *estruturais* em relação à canção. Mas a diferença também é fundamental para a Teoria dos Sistemas. Se, por um lado, a repetição é o processo que sedimenta as estruturas, a diferença é o que as instabiliza e modifica sua forma, fazendo-a crescer. “Assim, a Teoria dos Sistemas não começa sua fundamentação com uma unidade, ou com uma cosmologia que represente essa unidade, ou ainda com a categoria de ser, mas sim com a diferença” (LUHMANN, 2009, p. 81). Eis o seu pressuposto teórico. Parte-se disso para encontrar o sistema, que deve ser “a *diferença* resultante da *diferença* entre sistema e meio. O conceito de sistema aparece, na definição, duplicado no conceito de diferença” (LUHMANN, 2009, p. 81).

Em seguida, o autor aproxima o conceito de diferença com o de informação. Como vimos, o meio é diferente do sistema, por isso capaz de irritá-lo. Essa irritação aos poucos se transforma em informação para o sistema, que “atualiza as estruturas”, segundo Marcondes Filho (2004, p. 428). Por isso é preciso inverter o ponto de vista dependendo do que estou a expor aqui. Quando pretendo ver como a Música diferencia-se de si através do audiovisual, a primeira é sistema, o segundo meio. Agora, no que diz respeito ao problema de pesquisa, e ao modo como serão analisados os vídeos, a Música passa a ser meio, e o audiovisual, sistema, porque importa aqui mostrar como a Música interfere nos processos audiovisuais de modo a produzir imagens-música. E assim, de uma diferença em direção à outra, o sistema se modifica incessantemente.

Nos extremos do sistema, marginais ao meio, essa movimentação está ocorrendo. É exatamente aí que 4'33" de John Cage, por exemplo, se inscreve, provocando fortes irritações à Música. Neste limiar também estão os vídeos para *web* aqui estudados. Eles estão apresentando para a Música a possibilidade de ela se manifestar através de audiovisualidades. Este é um espaço de guerra, de embate, de jogo. Eis onde ocorre a autopoiesis, “determinação do estado posterior do sistema, a partir da limitação anterior à qual a operação chegou” (LUHMANN, 2009, p. 112).

No epicentro do sistema Musical encontram-se suas estruturas. Contam com as que mencionamos ao longo do texto, os elementos que se repetem nas músicas: introdução, verso,

refrão, solo, coda, modos gregos, escala temperada, padrões de afinação. Novas estruturas também ali se encontram, se sedimentando, num processo de *auto-organização*, que “significa construção de estruturas próprias dentro do sistema. Como os sistemas estão enclausurados em sua operação, eles não podem conter estruturas. Eles mesmos devem construí-las [...]” (LUHMANN, 2009, p. 112).

Se estamos entendendo os sistemas como organismos complexos, podemos pensar nos seus elementos constituintes como células que fazem a manutenção das estruturas. Essa manutenção só pode ser de ordem comunicativa, pois uma deve passar informação para outra ou para várias, de modo que se sedimente uma estrutura ou que se expulse um elemento que irrita o sistema. Para Maturana, “a linguagem não pode ser entendida como transmissão de algo, mas sim como uma supercoordenação da coordenação dos organismos” (apud LUHMANN, 2009, p. 294). A partir dessa proposição, Luhmann desenvolve a sua ideia de comunicação. Para ele, a comunicação só surge da síntese de três diferentes seleções, indissociáveis umas das outras: da seleção da informação, do ato de comunicar e do ato de entender ou não entender as duas primeiras. “Portanto, a comunicação acontece exclusivamente no momento em que se compreende a diferença entre *informação* e *ato de comunicar*” (idem, p. 297). Este entendimento da comunicação propõe que ela seja muito mais do que mera transmissão de informação, especialmente porque conta também como parte do processo o seu entendimento. Este terceiro passo tem a opção de refutar ou se apropriar da ideia, nos moldes da autopoiesis. A comunicação, portanto, é central para o entendimento da Teoria dos Sistemas, que enfatiza

a verdadeira *emergência* da comunicação. Não há propriamente transmissão de alguma coisa; mas sim uma redundância criada no sentido de que a comunicação inventa a sua própria memória, que pode ser evocada por diferentes pessoas, e de diferentes maneiras. (LUHMANN, 2009, p. 299)

Ela seria um sistema que se apropria exclusivamente de elementos de outros sistemas para configurar sua memória, enquanto, simultaneamente, permite a autopoiesis dos sistemas nos quais opera. É a comunicação que informa as estruturas dos sistemas, e é pela comunicação que os sistemas refutam irritações.

Para nós, ocidentais, onde a Música expressa-se mais propriamente é no campo do sistema tonal, que normatiza praticamente toda música que ouvimos na tevê e no rádio, que “abrange o arco histórico que vai do desenvolvimento da polifonia medieval ao atonalismo [...], e tem seu momento forte entre Bach e Wagner (ou Mahler), do barroco ao romantismo tardio, passando pelo estilo clássico” (WISNIK, 1999, p. 9-10). Embora no meio erudito já

venha se fazendo músicas a partir de outros sistemas como o serialismo há quase um século, para boa parte das pessoas leigas o Sistema Música se confunde com o Sistema Tonal – por isso, diz-se que nascemos com o ouvido tonal, o que nos permite reconhecer quando uma música esteja fora desse padrão. Isso é possível porque o Sistema da Música está em comunicação com nossa consciência. Observamos a música em questão sob os olhos da Música, mas julgamos em nossos termos: só o que nos parece tonal é considerado Musical. É uma dupla observação, carregada de entendimentos e desentendimentos. “Esse é o sentido que Luhmann dá à comunicação: a capacidade de sistemas observarem e de se observar sua observação. É nessa dupla operação que os sistemas vivem e sobrevivem” (MARCONDES FILHO, 2004, p. 432-3).

Finalmente, agora é possível pensar numa comunicação, ou também numa dupla operação de observação entre a Música e o audiovisual. Como foi mencionado, historicamente, as aplicações musicais em audiovisuais sempre ocorreram nos termos do audiovisual. Até mesmo nos videoclipes, que são entendidos como o que há de mais musical no meio audiovisual! Acontece que eles quase sempre foram produzidos *a partir* de uma música preconcebida, e montados de modo a produzir uma experiência muito mais visual, *cinematográfica*, do que musical, expandindo o sistema audiovisual somente nos seus aspectos visuais. Em raros momentos a Música sofreu irritação por um videoclipe. Por isso a discussão, quando se trata de arte, estética ou linguagem sempre foi voltada aos aspectos visuais do videoclipe. Em termos musicais, não há acontecimento propriamente, e não há nada de comunicacional, pois “Para a Teoria dos Sistemas, a função da comunicação reside em tornar provável o altamente improvável [...]” (LUHMANN, 2009, p. 306). No videoclipe, tudo que é musical é *provável*. A Música transmite e o audiovisual recebe passivamente, sem ser irritado. É um *comum acordo*, não uma comunicação. Não há, para a Música, um novo estado.

É isso que os vídeos para *web* mencionados no início do capítulo evidenciam. Naquele mix entre *Fatties at the grocery store rant* e *Battle Cry*, o que importa é a música que emerge do cruzamento de seus fluxos, e que só existe nessa simultaneidade, nesse processo. E mesmo nesse terreno surgem introdução, refrões, enfim, as repetições das estruturas da Música. Nesses casos, o audiovisual abre, portanto, muito mais frestas, se abre para acoplamentos estruturais com o meio musical.

Há uma frase do filósofo brasileiro Gerd Bornheim que é a chave de tudo o que vem se desenhando neste capítulo: “Pode-se dizer que, de certo modo, não é o compositor que faz a música; muito mais, a música se faz através do compositor” (2001, p. 143). Essa proposição

é a chave para formular qualquer problema de pesquisa que queira dar conta dos processos criativos em Música, e também legitima o uso da Teoria dos Sistemas de Luhmann no entendimento da Música tanto no processo composicional como nos audiovisuais musicais. É mais do que tempo de, talvez não romper, mas ao menos ignorar a distinção sujeito/objeto, pois que todas as nossas ações de observação sempre são observadas pelos sistemas.

### 4.3 DO NÃO-SENTIDO AO SENTIDO ARTICULADO

Antes de falar dos vídeos musicais para *web*, faz-se necessário discutir como a Música é capaz de se significar, de produzir sentido. Se estamos tratando especificamente de música tonal, a unidade mínima da linguagem musical é a nota, um som com altura melódica determinada (descontando os instrumentos de percussão). Podemos aproximá-la ao fonema na linguística, pois uma melodia, assim como uma palavra, se forma através da composição diacrônica de várias notas em sequência. Como podemos ver em Deleuze, “[...] *o sentido resulta sempre da combinação de elementos que não são eles próprios significantes*” (2005, p. 244). Os sons (musicais) são capazes de comunicar apenas quando combinados, quando ao formar uma melodia que reconhecemos pertencer a determinada escala ou estilo musical. Estamos acostumados a enxergar isso em uma partitura, que nos apresenta um gráfico em que, na vertical, vemos a altura melódica, e na horizontal, as notas no tempo. A partitura permite que se desenvolvam várias melodias simultaneamente, o que se chama de polifonia. Se elas entram temporalmente em momentos contrastantes de consonância e dissonância, elas adquirem uma característica contrapontística. Foi o que inspirou Eisenstein (2002) a desenvolver a teoria da montagem *polifônica*, em que o panorama visual deveria acompanhar as notas transcritas em partitura da trilha sonora com seus elementos (cor, contraste, movimento, enquadramento, etc), e fazer isso através do choque de elementos visuais e sonoros sincronizados. Esta ideia pretende pensar a atualização da Música no visual a partir dos elementos rítmicos, melódicos e harmônicos representados pelas notas musicais, encontrados visualmente nas partituras. Eisenstein pensava assim poder simular visualmente a linguagem musical, aquela que diz que determinado som é Dó, Ré ou Mi. Embora Bornheim não trate deste caso em seu texto sobre a linguagem musical, a citação abaixo se adequa muito bem como uma crítica ao método eisensteiniano:

Obviamente, esses sinais nada têm a ver com a linguagem musical, a música não está na partitura: música sempre houve, e a partitura surgiu bem mais tarde. A linguagem musical reside no som, ou melhor, no som enquanto transformado em tom, isto é, no som inserido numa escala, numa frase, num sistema sonoro. E essa organização sonora expressa algo, forma uma linguagem não redutível à linguagem conceitual. (2001, p. 136)

A partitura apresenta imagens possíveis da Música, portanto. E de todas as suas possíveis atualizações, é certo que ela será mais potente como som, pois é como tal que muitas vezes se explica o que ela é. Certamente ela não é tão somente som, mas é imprescindível que haja som, ou ao menos a ideia de que haja som (como quando imaginamos uma música), para que ela possa se manifestar. Para Bornheim, “[...] é só pensando o som que lograremos atingir o elemento propriamente musical da música” (2001, p. 140). Isso não deve implicar que, ao estudar audiovisuais musicais, devamos nos ater a falar de Música apenas em sua materialidade sonora, mas pensar suas atualizações em elementos a partir das virtualidades, de potencialidade sonora, e não a partir de outro atual, como Eisenstein fazia com a partitura<sup>57</sup>.

Essa potência de som é o que dá poderes sinestésicos à Música, o que torna possível a reconhecer, em potência, onde não há som, como na trilha visual, por exemplo, e até mesmo na partitura. Deve-se precisar que isso lhe dá a capacidade de se atualizar também em matérias outras que a sonora. No entanto, não implica obrigatoriamente que ela seja capaz de representar outras matérias:

Daí o acerto de Boris de Schloezer ao afirmar que a obra musical não é signo de qualquer coisa que a transcenda e se esforça por esposar, porque se assim fosse ficaria eternamente aquém de um ideal inexequível: ao contrário disso ela “se significa”, como diz Schloezer, e esse auto-significar-se permanece físico, embebido no carnal. (BORNHEIM, 2001, p. 142)

Uma nota isolada, assim como o fonema na linguística, não é capaz de produzir sentidos. Qualquer relação entre sons só pode ser feita a partir da arbitrariedade relacional proporcionada pela linguagem – isto é, quando deixam de ser sons puros! “O que torna a linguagem possível é o que separa os sons dos corpos e os organiza em proposições, torna-os livres para a função expressiva” (DELEUZE: 2007, p. 187). Se para Schloezer, a Música “se significa”, por outro lado esse mesmo processo de significação, dependendo da interpretação, pode destituir o papel do som como *som*, podendo então significar, expressar algo, representar mundos. Um dos melhores exemplos disso em Música é o madrigalismo, um estilo em que os

<sup>57</sup>Prendergast (1992, p. 223-226) critica a análise que Eisenstein faz de uma sequência de seu próprio filme, *Alexander Nevski*. O cineasta sugere que há uma correspondência completa entre imagem e música. Entre várias questões, Prendergast alerta, por exemplo, que a ideia de ritmo, que é bem clara na música, é metafórica nas artes plásticas.

sons utilizados expressam sentidos, como um glissando do grave para o agudo para significar que determinado personagem subiu um morro, ou batidas de tímpano e pratos representando uma tempestade.

Para nos opormos a isso, basta lembrar a visão (ou melhor, a audição) de Cage (1991): “[...] eu amo sons, exatamente como são. E não tenho a necessidade de que eles sejam mais do que são. Não quero que o som seja psicológico, nem que pretenda ser um balde ou presidente, ou que esteja apaixonado por outro som. Só quero que seja som”. Quando o compositor deu essa declaração, estava se referindo ao um dos objetivos de sua obra, que é o de desafiar o ouvinte a parar de significar os sons. E com isso fica claro por que é tão difícil perceber os sons como sons: é porque o conceito que temos da coisa vem à tona logo que é percebida, porque todo som tende a ser signo, mais ou menos nos termos da imagem acústica de Saussure (1993). “Como diz Bergson, não vamos dos sons às imagens e das imagens ao sentido: instalamo-nos logo 'de saída' em pleno sentido” (DELEUZE, 2007, p. 31). Essa condição inquietou muitos músicos modernos, como Cage, mas também como Igor Stravinski, que esquentou o debate na década de 1930 com suas crônicas:

Seu texto dizia que a música era "incapaz de *exprimir* seja o que for: sentimento, atitude, estado psicológico, fenômeno da natureza". A *expressão* jamais foi propriedade imanente da música. "Se a música parece exprimir alguma coisa, é apenas ilusão e não realidade (...), um elemento adicional que por convenção tácita e inveterada emprestamos, impusemos a ela, como uma etiqueta, um protocolo, uma roupa enfim que por hábito ou inconsciência acabamos por confundir com sua essência. (COELHO NETTO, 1995, p. 135)

No livro *O ouvido pensante* (1991), Schafer apresenta uma distinção radical entre linguagem e Música: “Linguagem é comunicação através de organizações simbólicas de fonemas chamados palavras. Música é comunicação através de organizações de sons e objetos sonoros. *Ergo*: Linguagem é som como sentido. Música é som como som” (1991, p. 239). A linguagem pode permear os eventos musicais, mas isso não quer dizer que a Música seja uma manifestação estritamente linguageira.

O debate é acalorado porque as leituras de cada autor apresentado aqui são muito distintas. Mas a que vamos adotar aqui é a de Bergson: quando ouvimos um som, primeiro compreendemos seu sentido, para só depois refletir sobre as características timbrísticas do mesmo.

Se a maior parte dos vídeos aqui estudados apresenta músicas tonais manifestadas na trilha sonora, como esses sentidos emergem? A construção ou destituição de sentido ocorre nos dois eixos do audiovisual (que é também o da música manifestada sonoramente): o

diacrônico e o sincrônico. O primeiro se dá no desenrolar do tempo linear, na sequencialidade, tanto da trilha visual quanto da trilha sonora. O segundo se dá no eixo horizontal, das articulações entre os elementos visuais e sonoros.

O primeiro eixo, considerando somente a trilha sonora, já foi abordado neste capítulo: implica em reconhecer que o sentido emerge da articulação entre sons, seja entre notas musicais (dentro do sistema tonal) que compõem melodias, seja entre fonemas que compõem signos verbais. Na trilha visual, é a montagem que vai exercer essa função de fazer com que planos se articulem para direcionar o espectador para um sentido específico. Notas musicais ou planos visuais, individualmente, só são capazes de referirem-se a si próprios nos vídeos musicais para web aqui estudados.

O método de produção de *samplers* audiovisuais, que já foi mencionado anteriormente, demonstra como a Música interfere nos métodos de mixagem dos vídeos, quando os usuários fatiam as sequências audiovisuais de vídeos (não-musicais) arquivados no *YouTube* em células que remetem a notas musicais isoladas em uma partitura (ver capítulo 5.3 para mais detalhes). Se não, os usuários optam também por recortar frases de modo a ressignificá-las, alterando a ordem. Essas frases, embora pertençam ao domínio da linguagem verbal, a sobrecodificação da Música incorpora-os ao convertê-los em imagem-música, imprimindo-lhes ritmicidade (*sampling* audiovisual), ou até melodizando-as (*auto-tuning*). No caso do já mencionado *MC Jeremias*, de um trecho do vídeo, em que se ouve o seguinte diálogo,

*GENIVALDO - Rapaz, tu bebesse demais, não foi?*

*JEREMIAS - Bebi, [?] diga ao delegado que eu sou cabra homem. Se eu pudesse eu matava mil!*

*GENIVALDO - Mas o quê, rapaz? Cachaça, cerveja...?*

*JEREMIAS - Eu mato até o delegado se ele for abusivo. Vai se foder todo mundo que eu já tô fodido.*

*GENIVALDO - Mas vem cá, como foi, tu tava bebendo onde?*

*JEREMIAS - Como foi? Eu tava bebendo no inferno! O Cão foi quem botou para nós beber! E você pode dizer ao delegado que ele for... Diga que eu já tô lascado, já tô fodido...*

*[...]*

*GENIVALDO - Canta uma musiquinha aí, vá!*

*JEREMIAS - Ei, ei, ei, é, sem você não viverei, todo o amor que eu te dei...*

*GENIVALDO - Todo o amor desse mundo... vá!*

*JEREMIAS - ...Sem você não viverei...*

Extraiu-se orações que viraram versos para o funk carioca:

*- Rapaz, tu bebesse demais, não foi?*

*- Bebi! (Bebi... bebi...)*

*Se eu pudesse eu matava mil  
que eu sou cabra homem  
Se eu pudesse eu matava (mil, mi-mil mi-mil)*

*Ei, ei, ei, ê  
tava bebendo no inferno  
Ei, ei, ei, ê  
Cabra safado tem que morrer  
Ei, ei, ei, ê  
tava bebendo no inferno  
Ei, ei, ei, ê  
O cão foi quem botou para nós beber!*

Vale lembrar, como já foi mencionado anteriormente em 1.1, que a trilha visual do vídeo original é cortada tal como a sonora, e o que temos em “*mil, mi-mil mi-mil*” no remix é uma “gagueira” provocada pela repetição do fonema, que nesse caso é composto por *samplers* audiovisuais, implicando também numa gagueira visual. Os gritos de Jeremias se transformam em versos similares aos proferidos por *MCs* de funk carioca.

O segundo eixo, das relações entre o visual e o sonoro, nos vídeos aqui analisados, implica basicamente em reconhecer de qual fonte os sons provém (poder ver o Jeremias proferindo os versos), ou com qual imagem visual elas se articulam. Como se trata de uma verticalização, o sentido emerge na simultaneidade de duas imagens, ao modo dos ideogramas chineses, bem como Eisenstein procurou associar ao cinema sonoro. Claro que vai muito além do que simplesmente sincronizar “a bota com o seu rangido”, como o próprio autor comenta (2002, p. 59). Há uma sincronização interna oculta, que deve ser operacionalizada a partir de uma linguagem comum a ambos, audiovisual e música: o movimento. Quando assistimos aos *spoofs* de *shreds* ou a um mashup que sincroniza os movimentos labiais e corporais do Seu Madruga com a canção *Florentina* do Tiririca, é através da correlação do movimento da trilha sonora com a visual que se faz entender que o ator Ramón Valdez, mesmo morto há quase trinta anos, está cantando *Florentina*, canção composta na década de noventa. Embora o espectador provavelmente saiba que se trata de um truque de edição, mesmo assim, essa sincronização é capaz de produzir, para a percepção do espectador, uma imagem-música (da ordem do possível) de Seu Madruga cantando *Florentina*. Uma imagem potencial dessa ordem só poderia ter sido produzida a partir de imagens técnicas.

Em termos de linguagem, também é possível compreender da seguinte forma: essa imagem-música (que pode ser entendida como um signo, um sentido), só emergiu da síntese de dois fluxos (Seu Madruga na trilha visual e *Florentina* na sonora) que não possuíam relação alguma. Cada um produzia sentido por conta própria, mas, ao serem unidos em



sincronia, eles paradoxalmente produziram sentido, no plano do *possível*, mas que não faz o menor sentido no plano do *realizado*.

Há, portanto, dois caminhos que os processos de composição dos vídeos adotam e que têm relação com o que está sendo debatido neste capítulo. O primeiro deles, com ênfase na diacronia, conta com três etapas: o usuário do *YouTube* seleciona um vídeo não-musical, como uma reportagem jornalística, por exemplo, em que reconhece um devir musical, baixa este vídeo para seu computador e abre o arquivo em um *software* de edição não-linear (esse vídeo tem sentidos construídos para um propósito específico); em seguida, ele fraciona o vídeo, selecionando momentos mais oportunos para que sejam convertidos em música (processo de *sampling* audiovisual, que destitui os sentidos originais do vídeo); selecionados esses momentos, eles são rearranjados pelo compositor, de modo a constituir uma música através de audiovisuais, processo que é regrado pelos termos da Música (estabelecendo novos sentidos, que provêm da articulação de elementos desprovidos do mesmo). Esse caminho, como será abordado no próximo capítulo, ocorre com mais frequência nos vídeos de *sampling* audiovisual e *auto-tunning*.

O segundo caminho tem ênfase maior nas sincronias, e ocorre com mais frequência nos vídeos de *Mashups* audiovisuais e *spoofs* de *shreds*. Nesses casos, o usuário baixa dois arquivos para seu computador, como, por exemplo, um arquivo de vídeo e outro de áudio, sendo que do primeiro são aproveitadas apenas as imagens visuais; a partir da trilha de áudio, que contém uma canção, ele repete os passos do primeiro caminho com as imagens visuais, mas as rearranja para sincronizá-las de algum modo (sincronização labial, movimentação, ritmo correlatos etc.) com a canção que ocupa a trilha sonora.

Pensar nesses eixos ajuda a compreender como os vídeos para *web* apresentam um novo panorama para a Música se manifestar, como veremos a seguir nas análises, quando se recorrerá a estes eixos não só para se reconhecer distinções, diálogos e sincronização entre trilha visual e sonora, mas também para ultrapassar os processos de composição musical e audiovisual em busca da imagem-música.

Estão estabelecidos, portanto, os procedimentos de análise da imagem-música em vídeos para *web*, a ser desenvolvida em três etapas. De acordo com o gráfico 3, apresentado no capítulo 2.4, se partirá do realizado, da observação empírica, para encontrar atuais e possíveis de imagem-música, e então reconhecer os processos de diferenciação que elas provocaram na Música:

1º) adota-se, através do pensamento sistêmico, a observação sistema/meio, compreendendo primeiro a Música como sistema e o audiovisual como meio: tal ponto de

vista permite, partindo do realizado (um vídeo musical para web, no caso), compreender como a Música diferencia-se de si através do audiovisual, produzindo atuais ou possíveis de imagem-música;

2º) em seguida, inverte-se o ponto de vista, partindo das imagens-música atuais e possíveis reconhecidas, entendendo o audiovisual como sistema e a Música como meio, para compreender como a Música interfere nos processos audiovisuais, permitindo que se descreva e compreenda as imagens-música identificadas;

3º) por fim, devolve-se as imagens-música ao virtual, compreendendo como elas movimentaram as estruturas da Música, forçando-a a criar novas informações.

Esses três passos são fundamentais para que seja possível o reconhecimento das imagens-música.

## 5 A IMAGEM-MÚSICA NOS VÍDEOS PARA WEB

Este último capítulo dá conta das análises dos vídeos musicais para *web*. Como foi exemplificado no capítulo 4.3, pretendo encontrar atualizações de imagem-música, manifestações de diferenciação de si, atuais e possíveis que fazem expandir o virtual da Música. Para tal propósito seguirei os procedimentos metodológicos descritos no capítulo anterior. Dependendo de como o objeto de análise se apresenta, os caminhos percorridos pela análise podem mudar, mas – na maior parte das vezes – se dá dessa forma: primeiro, compreende-se como os vídeos foram compostos, de que modo eles destituem o sentido de imagens de arquivo para então ressignificá-las em forma de vídeo musical; em seguida, compreendendo a Música como sistema e o audiovisual como meio, se reconhecerá acontecimentos nos vídeos em que o virtual da Música diferencia-se de si; por fim, invertendo esse ponto de vista, compreendendo o audiovisual como sistema e a Música como meio, percebe-se como ela irrita o audiovisual a ponto de se produzirem imagens-música. Estão divididos em práticas recorrentes, e é por isso que trabalho com muitos vídeos constituindo o corpus. No entanto, não pretendo analisar todos: embora já tenha mencionado vários ao longo dos capítulos anteriores, optei por eleger alguns que fossem mais interessantes por serem mais complexos e significativos para a análise. São os representantes onde a Música se manifesta, com maior potência, como produtora de novos problemas; vídeos nos quais a irritação que a Música exerce no audiovisual é convertida em informação, e que melhor definem as categorias que aqui apresento. Deve-se reparar também que estão organizados em distanciamento das práticas anteriores (cinema e videoclipe): os *mashups* audiovisuais guardam várias semelhanças com os processos de edição de videoclipes massivos, pois na maior parte das vezes eles são compostos de uma música pré-concebida na trilha sonora e de imagens visuais que são sincronizadas de acordo com essa canção; os vídeos de *sampling* audiovisual já são um pouco distintos, porque a música resulta da montagem: nesse caso, o *software* é utilizado como um tipo de instrumento musical; os *spoofs* de *shreds* invertem a lógica dos *mashups*, pois nesse caso a edição ocorre na trilha sonora, que é composta baseada no que comparece na trilha visual, provinda de uma performance musical de um músico famoso pré-concebida; os vídeos de *auto-tunning*, por sua vez, partem da mesma lógica dos *samplers* audiovisuais, interferindo na pós-produção na trilha sonora, convertendo depoimentos de entrevistados e falas de âncoras de jornal em melodias musicais; os vídeos que imaginam 4'33" já destoam completamente dos videoclipes, indo para o campo do erudito

e não apresentando uma música no formato canção; embora estes não possuam muita relação com os outros quatro, eles fecham a análise como problematizadores das outras categorias, como linha de fuga que apresenta outras possibilidades de diferenciação da Música em vídeos musicais para *web* que não puderam ser aqui abordadas ou que ainda não foram manifestadas, mas que são potentes.

As relações entre o visual e o sonoro continuam potentes, motivo pelo qual serão mencionadas diversas vezes, mas vale dizer que são observadas para apontar uma irritação da Música convertida em informação pelo audiovisual. O objetivo aqui é encontrar atualizações de imagens-música nos encontros entre imagens técnicas. De novo, é curioso como Flusser antecipou a situação na qual esta análise se insere: “Escrever, no passado, significava transformar imagens opacas em imagens transparentes para o mundo. Significará, no futuro, tornar transparentes as tecno-imagens opacas para os textos que estão escondendo” (2007, p. 149-150). O que se faz nas análises a seguir é justamente isso. O procedimento de análise trata de ultrapassar a opacidade das imagens técnicas, que procuram disfarçar seus modos de constituição, para encontrar a imagem-música. Isso será feito através do cruzamento da observação empírica com as teorias que foram apresentadas principalmente no capítulo 4.

## 5.1 MASHUPS AUDIOVISUAIS

Logo nos primeiros dois anos do *YouTube*, a primeira prática de *remixagem* encontrada, e com recorrência, é a que denomina-se aqui de *mashup* audiovisual, e que um dos exemplos mais sintomáticos é *System of a vila - Chaves Suey*. O autor do vídeo, PorCobain, montou trechos de cenas de alguns episódios do seriado *Chaves* em sincronia com a música *Chop Suey*, da banda *System of a down*.

Tanto no vídeo em questão quanto em vários outros que usam as mesmas imagens de *Chaves* para outras canções, vale considerar um respeito dos editores pela estrutura da canção. No primeiro verso, vemos uma sequência de imagens montadas de acordo com ritmos e melodias em destaque: os versos da canção, mesmo que sejam proferidas por um homem, são sincronizados para parecerem ser proferidas visualmente pela Chiquinha, personagem do seriado, enquanto que uma voz de fundo, sussurrada, é sincronizada com imagens em primeiro plano de outro personagem, Quico. Na repetição desse verso após o primeiro refrão, essa mesma sequência reaparece idêntica. Isso não é preguiça, mas é novamente a Música

exercendo o papel de sujeito que compõe as imagens visuais *através* do usuário. O mesmo ocorre nas repetições dos refrões.

No final da música, que sonoramente apresenta um momento apoteótico, redução do andamento e aumento de tensão, PorCobain as acompaha apresentando imagens das personagens de *Chaves* chorando ou gritando, balançando os braços e o corpo, mas as põe, através do software, em *slow motion*, prática raramente empregada no seriado, mas novamente provocada pela virtualidade da Música para forçar um entendimento de tensão máxima, de suspensão.

*Mashups* audiovisuais em geral provocam um choque contrapontístico, mas não nos termos de Eisenstein (2002), em que a oposição deveria se dar “no tempo”, melodicamente. No desenrolar das imagens, estas são compostas para se harmonizar das maneiras mais óbvias, até com corte *on the beat*. O choque advém, basicamente, do encontro de duas materialidades de universos muito diferentes: o *stone rock* da banda *System of a down* e o humor de *Chaves*. Essa potência geradora de improbabilidade que os softwares de edição não-linear de vídeo e áudio digital, aliados à biblioteca de imagens que é o *YouTube*, é o que vai desafiar seus usuários a produzirem montagens cada vez mais complexas e com esse caráter etéreo, fantasmagórico<sup>58</sup>, fantástico e improvável.

Pode-se dizer também que nos gestos dos personagens de *Chaves* há um devir-*heavy metal*, um pulso violento e descontrolado, e que serve como uma luva para uma canção do gênero, como podemos ver nos *frames* abaixo, retirados do *mashup* audiovisual em questão. No *frame* 9, Chaves gira uma matraca (que vemos apenas uma parte em quadro), se assemelhando a baixistas de *punk rock*, que seguram seus instrumentos na altura dos quadris e tocam corcundas e olhando para baixo. No 10, Quico grita com tamanha força que se pode ver as veias do pescoço saltando, tal como vocalistas de *stone rock*, assim como ocorre com Seu Madruga no *frame* 11, gesticulando enfaticamente. Essas relações aparentemente absurdas só foram possíveis de se constatar ao assistir ao *mashup* realizado, e pode-se, assim, encontrar uma imagem-música potencial na performance dos personagens do seriado, realizada no encontro destas imagens com as imagens da canção.

---

<sup>58</sup>Estes dois adjetivos estão relacionados com a fantasmagoria aparente do cinema mudo, em que os atores ali presentes visualmente eram desprovidos de suas próprias vozes, substituídas pelo ruído do projetor, enquanto que aqui são substituídas pelos ruídos do *Stone rock*.



Frame 9



Frame 10



Frame 11

Outro exemplo sintomático dessa prática é o vídeo *Seu Madruga cantando Florentina!*, que, assim como *Chaves Suey...*, denuncia ter sido produzido no *Windows Movie Maker* pois os créditos iniciais/finais foram desenvolvidos através de uma função automática de geração de caracteres do programa. Neste caso, cenas (então) raras do ator *Ramon Valdez* cantando acompanhado de uma banda são sincronizadas com a canção *Florentina* do músico e humorista brasileiro Tiririca. Novamente os gestos e movimentos do personagem na trilha visual vão se adequar aos movimentos sugeridos pela letra e sonoridades da trilha sonora.

Os gestos também se combinam potencialmente à canção da mesma forma que *Chaves Suey...*, mas dessa vez porque seus gestos se relacionam com algumas das ações sugeridas pelos versos de *Florentina* como quando Tiririca canta “[...] *senão eu boto esses teus dente prapara dentro*”, ao que visualmente é completado por uma imagem de Ramón Valdez cruzando os braços diante de seu rosto. E aí podemos anotar um dos primeiros elementos visuais em que a Música se faz presente: no gesto.

Uma outra relação audiovisual importante diz respeito ao sistema de *gestos* que acompanha toda interpretação musical. De fato, o discurso musical pressupõe, ao lado dos seus atributos mais profundamente acústicos, todo um sistema *kinésico*, entendendo-se como tal o conjunto dos elementos motores invocados pelo intérprete durante a performance. Em geral, todo esse trabalho gestual se perde no registro fonográfico e não é considerado na escrita formal da música: a notação. (MACHADO, 2009, p. 161)

Anos depois, em março de 2010, é publicado mais um *mashup* audiovisual envolvendo imagens de arquivo do seriado *Chaves*, dessa vez com a música *Bodies* da banda *Drowning Pool*<sup>59</sup>. Utilizando, na maior parte do tempo, as mesmas imagens que os outros dois vídeos mencionados, este vídeo produzido pelo usuário *H\$* reproduz as mesmas técnicas, só que com mais cuidados estéticos, aplicando efeitos de inversão e trepidação da imagem, típicos de videocliques de *nu metal* ou de *grunge* e *stone rock*, o que me faz trazer novamente Arlindo Machado para complementar o que digo, quando em *A televisão levada a sério* (2009), fala

<sup>59</sup>O nome do vídeo no YouTube é *El Chavo y su Banda - Drowning Pool feat Turma Do Chaves - Bodies in the Vila - -* e se encontra no seguinte endereço: <<http://www.youtube.com/watch?v=s-gjdeWj5b0>>. Acesso em: 24 set. 2011.

sobre videocliques da década de 1990 que usavam sujeira nas fitas, ranhuras na película e outros efeitos de deterioração como estética:

[...] é o abandono ou a rejeição total das regras do “bem fazer” herdadas da publicidade e do cinema comercial. O que vale agora é a energia que se imprime ao fluxo audiovisual, a fúria desconstrutiva e libidinosa que sacode e dissolve as formas bem definidas impostas pelo aparato técnico. [...] agora presenciamos o retorno a um primitivismo deliberado, à imagem “suja”, mal iluminada, mal ajustada, mal focada e granulada, o corte, a rebarba, a câmera sem estabilidade e sacudida por verdadeiros terremotos, todas as regras mandadas para o vinagre e todo o visível reduzido a manchas disformes, deselegantes, gritantes, inquietantes. (MACHADO, 2009, p. 177)

Em nenhum momento o seriado *Chaves* foi assim. No entanto, quando os usuários do *YouTube* foram em busca de suas velhas fitas *VHS* com programas gravados para digitalizá-los, as imagens naturalmente foram se deteriorando, e então é comum ver, nesses *mashups*, a imagem visual sofrendo *drop outs*, que é quando ocorre uma falha na leitura dos dados da fita (a característica são aqueles chiados horizontais na base da tela), ou ficando preto e branco do nada (falha de decodificação entre os padrões PAL-M e NTSC. Ver *frames* 12 a 15). Além disso, a imagem em *VHS* é muito problemática para ser digitalizada, o que acaba por evidenciar ainda mais aos olhos a compressão *MPEG* (formato de arquivo digital de baixa definição). Volto à questão da compressão para lembrar do seguinte: tanto a compressão de áudio quanto a de vídeo foram desenvolvidas para serem utilizadas após a finalização de um produto de alta definição – são arquivos, a princípio, apenas de visualização e audição. Mas não demorou muito para que os softwares de edição não-linear de áudio e vídeo permitissem que se pudessem utilizar tais formatos como material de edição. E é aqui que reside uma questão fundamental dos vídeos que estou analisando: boa parte deles são compostos de trechos de outros vídeos que já estavam no *YouTube*, que já haviam sido comprimidos, portanto. Devolvê-los à edição, para, então, comprimi-los novamente produz, sem dúvida, várias marcas, cicatrizes de uma manipulação excessiva, e que acabam por definir parte da estética dessas imagens, que poderíamos comparar à estética *grunge*, suja, maltrapilha, visceral e de baixa definição. É nesse momento que, acidente ou coincidência, as imagens digitalizadas de *Chaves* casam muito bem, agora em termos de textura, com a paisagem sonora de canções ruidosas de pós-rock. Cabe aqui comparar a textura, nitidez, tons de cor, objetos de cena e até vestimentas dos *frames* 12 a 14 (retirados de um intervalo de 3 segundos do vídeo *System of a vila - Chaves Suey*) com o *Frame* 15, retirado do videoclipe *Smells Like Teen Spirit* (Samuel Bayer, 1991), que possivelmente é o que melhor capturou o *zeitgeist* do período *grunge* do início da década de 1990.



Frame 12



Frame 13



Frame 14



Frame 15

Por fim, alerto que essa prática ainda está muito próxima da do videoclipe televisivo, de lançar uma música pré-gravada na trilha sonora e a partir dela compor a trilha visual. Mas é preciso lembrar a importância dessas primeiras práticas de interferência em imagens de arquivo, pois foram elas que iniciaram no *YouTube* tal cultura. Elas ainda não se alinham teoricamente com o que estou propondo como hipótese, mas já começam a ensaiar tal processo ao encontrar potências musicais no seriado *Chaves*, por isso sua importância neste projeto. Importante também lembrar que não só há uma grande quantidade de vídeos que fazem *mashups* de cenas desse seriado com canções contemporâneas, como também se produziu tal encontro com outros seriados e programas de TV. Ele também difere do videoclipe massivo porque não só não mostra os músicos que seriam os compositores, como tampouco apresenta pessoas fazendo *playback*. É diferente do que o diretor Johan Renck fez para o videoclipe *Crystal* da banda New Order (2001): inseriu uma banda fictícia no panorama visual. Nenhum dos músicos na tela integra oficialmente a banda. Já no caso desses *mashups* audiovisuais, o que eles fazem é simular a possibilidade de que personagens de um seriado da década de 1970 fossem capazes de executar uma canção que foi composta mais de trinta anos depois, fazendo, com esse encontro, que as imagens visuais manifestem suas potencialidades musicais.



## 5.2 SAMPLING AUDIOVISUAL

A prática de *sampling* audiovisual já é um passo a frente aos *mashups* (e um passo mais distante dos videoclipes massivos), e nem sempre utilizam imagens de arquivo do *YouTube*. Já expliquei ao longo do trabalho o que compreendo como *sampling* audiovisual, e mencionei como exemplos *Amateur*, *MC Jeremias* e o *Funk da menina pastora*. Quero, agora, me centrar em um caso bem específico e mais complexo: o dos denominados *VideoSongs*. Inspirado no já mencionado *Amateur*, o estadunidense Jack Conte resolveu transformar a ideia de Lasse Gjersten em uma escola. Identificando o *VideoSong* como uma “nova mídia”, Conte cria seus dogmas, conforme aparece na descrição de todos seus vídeos hospedados no *YouTube*: 1) O que você vê é o que você ouve (não há *playback* em instrumentos ou vozes); 2) Se você está ouvindo, em algum momento você irá ver (não há sons escondidos). É uma proposta bem diferente da dos videoclipes tradicionais, onde o que reina são os *playbacks*. É como se pudéssemos ver o processo de gravação das canções em multipista.

Com isso, Conte acabou criando uma verdadeira escola: hoje são vários os músicos que partilham da estética em seus vídeos no *YouTube*, especialmente nos moldes do *Frame 16*, partindo a tela em quatro molduras, cada uma exibindo um músico tocando um instrumento.



Frame 16

Enquanto Lasse Gjersten, que é um caso primeiro, editou trilha sonora e visual simultaneamente e diretamente em software de edição não-linear de vídeo, os *VideoSongs* de Jack Conte e sua banda Pomplamoose são compostos de modo diferente. Ele grava o áudio de um instrumento por vez, ao mesmo tempo em que se registra em vídeo, como podemos ver no *Frame 17*, extraído do *VideoSong Beat the horse*<sup>60</sup>, em que ele microfona uma meia-lua, enquanto ouve os instrumentos pré-gravados pelo headfone. Para cada trilha de áudio, ele tem uma trilha de vídeo respectiva.

<sup>60</sup>Disponível em: <<http://www.youtube.com/watch?v=8oJgqbgvInk>>. Acesso em 13 jan. 2012.



Frame 17

Assim que gravados todos os instrumentos, ele mixa o áudio, aplicando efeitos sonoros em cada instrumento e regulando o volume de cada trilha, e masteriza o som, fundindo as trilhas individuais em uma só. Em seguida, ele importa o áudio da canção para o software de vídeo, onde ela irá ser sincronizada com as tomadas de vídeo.

Como às vezes chegam a mais de uma dúzia de instrumentos, Conte não pode colocar todas as imagens de cada instrumento simultaneamente na tela, ao contrário de Gjesten, que usou apenas piano e bateria. Como um controlador de *switch* televisivo, Conte decide qual imagem deverá aparecer e em qual momento. Normalmente ele opta pelo instrumento que está em destaque em cada momento da música. Essa prática se opõe, a princípio, à ideia de *sampling* audiovisual, uma vez que oculta da trilha visual algumas imagens de instrumentos que estavam em sincronia com a imagem sonora desse instrumento na trilha sonora, mas, como explicarei a seguir, esse fundamento dos *VideoSongs* acaba por deixar ainda mais complexa a ideia de *sampling* audiovisual.

Quando os instrumentos fazem sons repetidos, ele prefere usar sempre o mesmo *sampler* audiovisual, como na sequência de *frames* abaixo (18 a 25), extraído dos três primeiros segundos de *Beat the horse*, onde podemos ver *samplers* audiovisuais de peças de bateria, tocadas em momentos diferentes, mas depois ordenadas de modo a compor um ritmo: bumbo-prato-caixa-prato-bumbo-prato-caixa-prato... (a caixa, como se pode ver nos *frames* 30 e 34, divide a tela com um bater de palmas). Aqui já temos algumas variações se comparado com *Amateur*: ao invés de todos os *samplers* audiovisuais serem capturados visualmente de um único enquadramento, cada unidade da bateria foi capturada por um ângulo diferente. A sequência usa apenas três *samplers* audiovisuais (com o *sampler* do prato intercalando bumbo e caixa), e, ao ser colocado em *loop*, constitui o ritmo de toda a canção, por onde os outros instrumentos a serem gravados irão se guiar (caso da meia-lua citada anteriormente), para poderem soar em sincronia.



Frame 18



Frame 19



Frame 20



Frame 21



Frame 22



Frame 23



Frame 24



Frame 25

Mas eu não estava satisfeito com a leitura que essa sequência de *frames* me apresentou. Eles não apresentam o que ocorre na trilha sonora, e pareciam esconder algo do processo de composição. Avancei para uma segunda análise, feita através do software *Final Cut* e editada no *Photoshop*. Com o *Final Cut*, foi possível apresentar visualmente o tempo de duração de cada plano visual (gráfico 4, primeira linha) e de cada evento sonoro (gráfico 5, última linha). Através do *Photoshop*, manipulei as imagens que aparecem na trilha visual para simular o exato momento em que há uma troca de plano, pois elas não aparecem desse modo no *Final Cut*. No topo, as marcações de tempo referem-se ao momento em que há uma troca visual de *sampler* na trilha visual, sendo que o primeiro dígito refere-se aos segundos, e o segundo, aos *frames* que compõem cada segundo (a cada trinta *frames*, um segundo se passa). Logo abaixo, no gráfico 5, está reproduzido o gráfico 4, mas com a adição de cores que representam cada instrumento. Os números indicam o tempo, em *frames* por segundo, que cada instrumento dura na trilha visual e na sonora. O gráfico mostra que, embora nos soe sincronizadas, não há uma correspondência absoluta entre as trilhas. Vale notar que não foi possível fazer uma marcação exata do início do ataque de cada instrumento da trilha sonora, pois ela não é composta através de *frames* como a trilha visual:

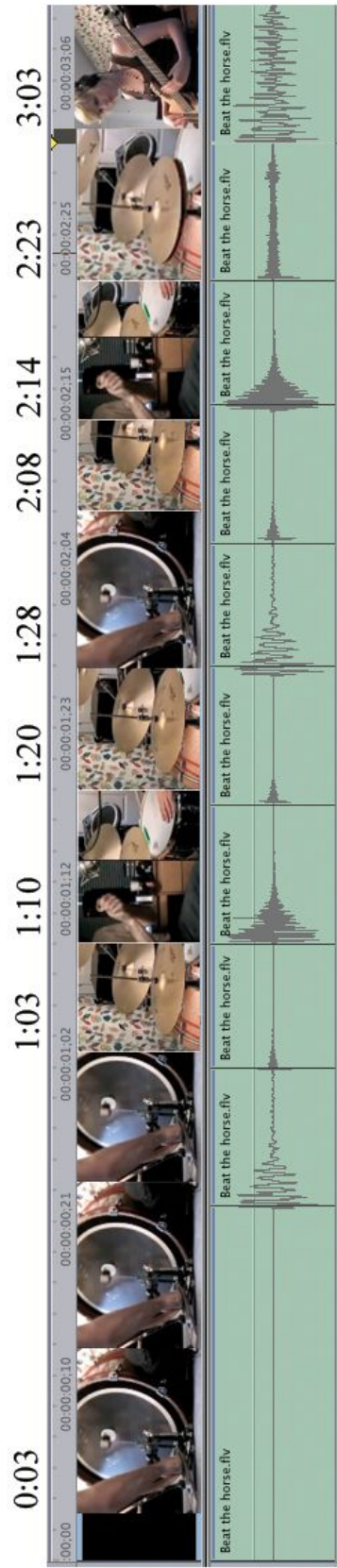


Gráfico 4



Gráfico 5

Lendo esse gráfico, a ideia de *sampler* ficaria menos potente porque, por mais que o visual e o sonoro tenham sido capturados ao mesmo tempo, o diretor decide as durações visuais, porque ele mixa em separado a trilha sonora, em *software* de produção de áudio digital, e depois ressincroniza os planos visuais um por um em *software* de edição não-linear. Parece, a princípio, que para haver uma melhor sincronização, os tempos de duração do visual *não* devem respeitar o sonoro (ora, o tempo de leitura do olho não é o mesmo do ouvido). A questão dos desencaixes tem resposta muito simples: embora tanto o áudio quanto o vídeo digital se projetem em fluxo (o primeiro vibrando alto faltantes e o segundo atualizando os *pixels* da tela, um por um), e não imagem estática após imagem estática como no cinema, o vídeo é montado em *frames*, grupos de pixels que formam uma imagem completa no panorama visual, e que duram trinta avos de segundo. Portanto, só será possível um encaixe perfeito de músicas que tenham um número de batidas por minuto que sejam múltiplas de trinta. De qualquer forma, essa velocidade é tão alta que nossa percepção não é capaz de encontrar tais dessincronias. Essa análise serve para comprovar que: 1) a análise anterior, a sequência de *frames*, dá mais conta de entender a imagem-música que a do gráfico, pois simula melhor como nossa percepção compreende o fluxo do vídeo; 2) que de fato a inserção das imagens visuais é manual, efetuada após a mixagem sonora; 3) que a noção de sincronia entre o visual e o sonoro é convencionalizada.

Interrompo neste momento a análise dos vídeos da Pomplamoose para me deter nesta terceira constatação. Recordei da época que editava videoclipes para minha banda, e lembrei que às vezes, para parecer mais natural a imagem do baterista atacando um prato, tinha de sincronizar o início desse evento sonoro às vezes um *frame* antes da baqueta efetivamente acertar o prato na trilha visual, às vezes um *frame* depois. Só percebi que isso era natural que ocorresse porque gravávamos as canções em software de edição de som, usando um metrônomo. Se uma canção tivesse 135 batidas por minuto, por exemplo, era óbvio que os *frames* não se encaixariam com perfeição aos ataques. Mas essa sincronia falsificada não ocorre apenas no audiovisual:

só ouvimos os sons (no sentido em que os reconhecemos) pouco depois de os termos percebido. Bata as palmas de forma breve e seca e ouça o barulho que fazem: a escuta – de facto, a apreensão sintética de um pequeno fragmento, guardado na memória, da história sonora – seguirá o acontecimento de muito perto, mas não lhe é totalmente simultânea. (CHION, 2008, p. 18)

Esta constatação Michel Chion devolve potência à ideia de *sampler* audiovisual, pois, mesmo que houvesse uma sincronia perfeita nas menores unidades do audiovisual (os

*frames*), nossa percepção não seria e nem é capaz de processar informações nessa velocidade de atualização. É nossa percepção que “decide” que evento visual pode e/ou deve entrar em sincronia com um evento sonoro que tenha ocorrido aproximadamente ao mesmo tempo, tanto que no cinema, se mostrarmos na trilha visual um cachorro latindo, mas na sonora trocar por um miado, compreenderemos esta imagem sonora do miado como provido da suposta fonte visual. Como evidência disso, podemos também lembrar de um vídeo fora de sincronia. Se assistirmos uma pessoa falando em primeiro plano, mesmo que haja dessincronia de um segundo, ainda reconheceremos que a voz que ouvimos provém da pessoa que aparece na tela. Nossa percepção é que faz a sincronia, independente do tamanho de defasagem entre o evento visual e o sonoro. Tratando especificamente do audiovisual, a sincronia é uma imagem técnica: ela carrega até hoje a magicidade, a fantasia de sonorização das imagens visuais que agregou nos primeiros filmes sonoros: ela é falsificante. Sabemos que diferentes aparelhos captam luz e sons, e que diferentes aparelhos os reproduzem, mas, novamente, como diz Bergson (apud DELEUZE, 2007, p. 31), nos colocamos de saída no sentido, então primeiro vamos perceber som e visual como possivelmente providos de uma mesma matéria. Não há um único fluxo em que tudo está em sincronia, mas uma multiplicidade de fluxos diferentes que se sobrepõem e que a nossa percepção trata de encontrar os pontos de síncrise<sup>61</sup>.

As constatações acima não só resolvem questões até então pendentes da relação entre o visual e o sonoro, como colaborarão mais adiante para o entendimento de outros vídeos aqui abordados, já mais voltados para a compreensão de outro ponto de vista, da relação entre Música e audiovisual.

Volto, agora, a analisar os vídeos da Pomplamoose.

Para enfatizar as relações dos *samplers* audiovisuais que fazem emergir canções, a maior parte do tempo do vídeo é composta por panoramas visuais com duas ou mais molduras contendo imagens de diferentes instrumentos, como mostra o *Frame 26* (extraído do *VideoSong My Favorite Things*<sup>62</sup>), em que Jack Conte, no violão, divide a tela com sua parceira Nataly Dawn replicada em outras três molduras, fazendo vozes em alturas diferentes em cada uma. Não há o tempo todo imagens dos instrumentos que são tocados como em *Amateur*, pois, ao contrário deste que tem somente bateria e piano, os vídeos de Jack Conte contam com vários instrumentos, mas ainda assim procura-se evidenciá-los em vídeo, às vezes até com o panorama sendo partido em nove (*Frame 27*). A complexidade visual, tanto na quantidade de molduras quando na velocidade maior de troca de *samplers*, acompanha os

<sup>61</sup>Para Chion (2008, p. 54), síncrise é a soma de sincronia e síntese, momentos específicos em que elementos em relação vertical entre dois fluxos se tornam mais brilhante do que as demais verticalizações.

<sup>62</sup>Disponível em: <<http://www.youtube.com/watch?v=OvYZMqOffQE>>. Acesso em 13 jan. 2012.

momentos de maior tensão musical: refrões, solos, codas.



Frame 26



Frame 27

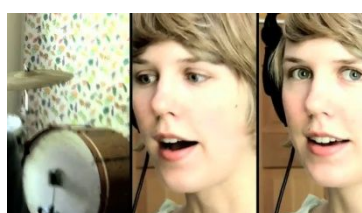
É preciso lembrar que, tanto em *Amateur* como nos *VideoSongs*, a música foi idealizada antes de ser registrada. Pode-se perceber isto, no primeiro citado, quando Gjersten toca algumas notas combinadas (acordes) no piano. Ele provavelmente compôs algum tipo de partitura para poder lançar posteriormente notas e acordes tocados em ordem correta no vídeo, senão teria que tocar o piano de todas as maneiras possíveis (que são infinitas) para poder compor livremente no *software* de edição – algo impossível de ser realizado. Nos *VideoSongs* de Jack Conte isso fica mais claro pelo fato de serem canções: os planos de pessoas cantando são poucas vezes *sampleados*, como ocorre em *MC Jeremias*. Mas a diferença essencial destes vídeos para o que até então se entendia como *videoclipe* é que essa pré-idealização pensa a Música como meio e o audiovisual como sistema desde o início, da necessidade de que tudo que se ouve também possa ser visto. E mesmo que deixem claro que a canção foi pensada antes da gravação do vídeo, ela só pôde ser ouvida após sua construção, realizada pela edição audiovisual. Este é irritado pela Música, mas apreende essa imposição em seus termos: por mais que a Pomplamoose quisesse apresentar os instrumentos visualmente o tempo todo, o audiovisual não permite que isso ocorra, por causa da tela de vídeo do *YouTube*, tão ou mais limitada do que a da televisão em quantidade de *pixels*.

Como dito anteriormente, o que está na trilha sonora foi capturado *em conjunto* à trilha visual, e não anteriormente, como ocorre em um videoclipe tradicional. Ela nasce *do* audiovisual, só foi possível porque o audiovisual se comportou de um modo nunca antes feito pelos videoclipes, onde tradicionalmente a canção utilizada provém de um disco lançado anteriormente.

Dessa descontinuidade sincronizada e replicada, emergem as canções que primam – surpresa – pelo fluxo contínuo de tensões e relaxamentos da canção popular. No *VideoSong*

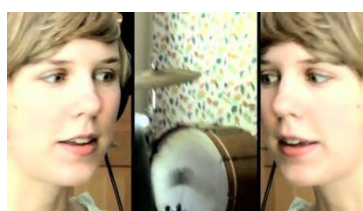
*Another Day*<sup>63</sup>, por exemplo, parte do refrão constitui-se da frase *I am yours*, cantada por uma dobra de voz por Nataly Dawn em *staccato* e acompanhada de um bumbo. Todos estes elementos podem ser vistos no vídeo, e a ênfase no *staccato* se dá pelo jogo de trocas de molduras entre as vozes e o bumbo (*frames* 28 a 30). Perguntas ficam no ar: os planos duram enquanto a nota dura, ou são cortados para que a frase musical possa durar? Algumas imagens, que duram menos de um segundo, ao serem montadas em sequência, imitam uma frase musical? Seriam essas durações audiovisuais na verdade *durações musicais*? São música em si ou representações musicais?

“I”



Frame 28

“am”



Frame 29

“yours”



Frame 30

Os *VideoSongs*, assim como *Amateur*, compõem um novo panorama para a relação entre Música e audiovisual. São o locus privilegiado da experimentação com ou contra o aparelho referido por Flusser. Então – ao invés de a música ser aquele elemento constituinte que ocupa apenas a trilha sonora do audiovisual, enquanto a trilha imagética é ocupada pela imagem do *superstar*, tal como nos videoclipes massivos – no *VideoSong* ela emerge como uma potencialidade em todos os elementos audiovisuais, tanto na trilha sonora como na visual. Ela se atualiza como imagem-música através dos sons, da montagem videográfica, dos entrefluxos (irritando o audiovisual, sendo este sistema e ela, meio), bem como comparece

<sup>63</sup>Disponível em: <<http://www.youtube.com/watch?v=Uolz7V12evc>>. Acesso em 13 jan. 2012.



como *virtualidade estruturante* (agora como sistema, e o processo composicional da Pomplamoose como meio).

No entanto, como não somos capazes de perceber o todo dessa estrutura, porque ela é dotada de certa invisibilidade, nós a percebemos quando se atualiza, isto é, quando algo ou parte de sua totalidade se imprime em determinada materialidade, como os *VideoSongs* referidos, cuja existência instaura mudanças também nos diagramas do virtual. Assim, Deleuze aproxima o conceito de virtual com o de estrutura, justamente para dar esse caráter de movimento incessante de atualização:

[...] o virtual tem uma realidade que lhe é própria, mas que não se confunde com nenhuma realidade atual, com nenhuma realidade presente ou passada; ele tem uma idealidade que lhe é própria, mas que não se confunde com nenhuma imagem possível, com nenhuma ideia abstrata. Da estrutura, diremos: *real sem ser atual, ideal sem ser abstrata*. (2005, p. 250)

Seriam, então, os *VideoSongs* atualizações da Música neste espaço anteriormente dominado pelos videoclipes? Tais vídeos produzidos para a *web* estariam criando novas zonas problemáticas tanto para as pesquisas quanto para as práticas audiovisuais? Deleuze nos apresenta um atalho: “A questão: há estrutura em qualquer domínio? deve, pois, ser assim precisada: podemos, neste ou naquele domínio, extrair elementos simbólicos, relações diferenciais e pontos singulares que lhes são próprios?” (idem, p. 222). Sem dúvida conseguimos extrair tais coisas de um *VideoSong*, mas como fazer isso pensando na Música como o que a estrutura<sup>64</sup>?

Assim, os *VideoSongs*, que são o pleno sentido desta pós-história referida por Flusser, podem para além e aquém de sua teoria, ser também compreendidos como uma resposta atual à Música, aqui considerada como estrutura nos termos deleuzeanos ou virtualidade nos termos bergsonianos. Nos vídeos compostos pela dupla Pomplamoose, as imagens visuais remetem à Música, raramente fazendo referência à letra da canção. Nelas, a Música manifesta, nos seus termos, o desejo de Cage, porque seus elementos significantes estão enfraquecidos ou ausentes de sentido, mesmo quando somados ou justapostos.

Se fizermos a leitura tradicional, separando a trilha visual da trilha sonora, veremos que o modo como se faz a edição das tomadas audiovisuais produz sentido em sua manifestação sonora (criam uma canção), enquanto que isso não ocorre em sua manifestação visual: o que se produz visualmente é um não-sentido, a descaracterização do fluxo musical pela descontinuidade e repetição de elementos visuais. Estariam os *VideoSongs* criando uma oposição à teoria de Eisenstein (2002, p. 14), que defende que “*dois pedaços de filme de*

<sup>64</sup> Estrutura é aqui compreendida de acordo com Gilles Deleuze.

*qualquer tipo, colocados juntos, inevitavelmente criam um novo conceito, uma nova qualidade, que surge da justaposição”?*

Minha resposta para o modo de produção de sentido nos *VideoSongs* é que não são mais dois pedaços de filme (no caso do *VideoSong*, dois *samplers* audiovisuais) que vão criar um novo conceito. A justaposição de um bumbo seguido de um prato não constitui um ritmo musical. Apenas quando esses elementos são postos em repetição, como o que ocorre nos *frames* 28 a 35, é que temos produção de sentido, porque dali emergirá uma musicalidade. É o visual se adequando às exigências Musicais. A relação que importa não é mais entre uma imagem visual e outra, tampouco entre o visual e o sonoro, mas sim entre Música e audiovisual, mais especificamente da capacidade de este último criar representações audiovisuais da Música. A este procedimento estruturante denomino imagem-música. A imagem-música põe em relação não dois termos – a imagem e a música –, mas estruturalidades de diferentes níveis (virtuais, atuais) cujos procedimentos tendem a gerar imagens capazes, por sua vez, de problematizar inclusive as relações que os fundaram. Isto é o que ocorre com os *VideoSongs* e com suas relações com o videoclipe, por um lado, e, por outro, com os aparelhos técnicos (no sentido flusseriano) que lhes enformam.

A imagem-música nos *VideoSongs* atualiza-se através da diferença e da repetição de *samplers* audiovisuais. A repetição foi exemplificada neste artigo com os *loops* de bateria sob os quais a canção *Beat The Horse* é constituída. A esse modo de repetição circular Deleuze e Guattari dão o nome de ritornelo: “precisamos criar um território, um ritornelo, para então deixar entrar um pouco de caos e podermos escapar. Para mais na frente criarmos novamente um outro ritornelo, e assim por diante...” (1997 apud NASCIMENTO, 2001, p. 67). Os ritornelos criam a base dos blocos da estrutura musical. No caso da canção popular, podemos entender o verso e o refrão como dois tempos antagônicos com ritornelos próprios que diferem um do outro, e cuja alternância repetitiva gera tensões e relaxamentos de uma música tonal.

A diferença ocorre quando se utiliza outro *loop* em complementação ou substituição ao anterior. Logo após o *loop* de bateria já mencionado (*frames* 18 a 25), um plano de um baixo sendo tocado substitui a imagem do bumbo, mas ainda se ouve o instrumento de percussão. A imagem é substituída, enquanto o som é complementado. Em vários momentos dos vídeos do Pomplamoose, alguns *loops* e outros tipos de imagens repetitivas são substituídas na trilha visual mesmo que permaneçam se repetindo na trilha sonora, parte pela dificuldade de mostrar todas as trilhas visuais num só panorama (a definição dos vídeos para a *web* é igual ou mais baixa que a da televisão), parte porque isso embora funcione para a manifestação

sonora da música, nem sempre funciona na manifestação visual, o que torna ainda mais complexo o problema de pesquisa que estou apresentando. Por outro lado, essas imagens visuais que somem são capazes de se sobrepor às imagens em fluxo como imagem-lembrança. Uma vez apresentadas, somos capazes de deduzir sua função estruturante mesmo em sua ausência. A Música opera da mesma forma, tanto que é por isso que tendencialmente canções populares começam e terminam com a nota que dá o tom: mesmo que ela não seja tocada novamente até o fim, ela define o centro de gravidade por ter sido tocada primeiro. A diferença emerge da tensão de elementos territorializados, do lugar-comum, do ritornelo. Só se produz diferença partindo da repetição, pois, somente como memória, como virtualidade, que a diferença será capaz de tensioná-la.

Diferença e repetição são dois vetores fundamentais para que se pensem com categorias novas os vídeos produzidos para a *web*: os *VideoSongs*. Se, por um lado, a repetição cria hábitos tanto de composição quanto de produção teórica (o videoclipe e suas teorias são um exemplo evidente de tal procedimento), a diferença é o procedimento que instabiliza tanto produções quanto teorias, modificando-as e as fazendo crescer em múltiplas semioses. As condições de crescimento de tais semioses não estão apartadas de seus processos de produção. É provável que no sistema tradicional da indústria fonográfica o *VideoSong* tivesse dificuldade de expressão, pois neste território músicos e produtores de vídeo vivem separados, com funções bem específicas e separadas no tempo e no espaço, como ocorreu com Godley e Creme e seu *Mundo Vídeo*. Na *web*, ao contrário, a produção deixa de ser necessariamente em série e especializada: a imagem do capitalismo da pós-história é a das sincronicidades que têm gerado os *VideoSongs* a partir de aparelhos técnicos cujos programas imaginam mundos por vir a partir de seus próprios termos.

Assim o mérito do *VideoSong* parece estar em ter conseguido desterritorializar a Música no terreno dos audiovisuais musicais; não produzir uma música diferente, mas produzir diferença na Música, usando o audiovisual como um instrumento musical capaz de imaginar Música, de apresentá-la teoricamente através de imagem-música. É na busca por um novo ritornelo que a diferença dará movimento ao virtual. Nessa reterritorialização que o *VideoSong* proporciona à Música não há uma linguagem estabelecida, mas um território movediço, onde a Música é quem protagoniza a produção de sentido.

### 5.3 SPOOFS DE SHREDS

Talvez a melhor tradução para *shred* seja “fritada”. O termo em inglês é utilizado para descrever solos de guitarra de altíssima velocidade, o que também é conhecido por *virtuose* ou, no Brasil, pela gíria “fritar”. No *YouTube* constam inúmeros vídeos de guitarristas famosos fritando (*shredding*). Eis que Santeri Ojala, ou StSanders, seu nome de usuário no *YouTube*, baixa estes vídeos, retira a trilha sonora, e no lugar, insere sua própria performance sonora de fritada, propositadamente muito mal executada, mas tentando imitar nos sons os movimentos aparentes dos guitarristas. O que ele faz é, ao ver o guitarrista solando na parte mais aguda da guitarra, próxima ao corpo dela, dubla com sua própria guitarra na mesma região, mas tocando notas erradas, fora do tom, ou tocando em falso a palheta. Aplicando ainda imagens sonoras de arquivo de plateia e reverberações típicas de execuções ao vivo, StSanders consegue promover um choque entre o visual e o sonoro inverso ao dos *mashups* audiovisuais e dos videoclipes massivos: ao invés de mimetizar visualmente eventos sonoros, os *shreds* tratam de traduzir sonoramente imagens visuais, especialmente os gestos e movimentos labiais de músicos em performance. À primeira vista, trata-se de um caso simples de relações entre o visual e o sonoro, mas o ponto de vista que adotamos permite perceber o choque que a Música exerce no audiovisual.

Não só StSanders, mas vários outros usuários do *YouTube* nele inspirados produziram incontáveis *spoofs* de *shreds*. Grandes guitarristas como Slash<sup>65</sup>, Eddie Van Halen<sup>66</sup>, Carlos Santana<sup>67</sup> e Steve Vai<sup>68</sup> caíram nas mãos dos falsificadores de solos.

Um caso curioso merece atenção: um *spoof* de uma performance de Frank Zappa<sup>69</sup>, produzido pelo usuário BluesMK. Conhecido por seu estilo ousado, influenciado pela vanguarda erudita, é possível que Zappa soe, para os ouvidos acostumados ao tonalismo, como um péssimo guitarrista. Ao assistir ao vídeo original, que também se encontra no *YouTube*<sup>70</sup>, repara-se que a estranheza é causada principalmente pela mistura de estilos e escalas musicais, o uso recorrente e intenso de alavanca (mecanismo de mola instalado no corpo da guitarra que solta a tensão das cordas ao ser pressionado, utilizado para produzir

<sup>65</sup>Disponível em: <<http://www.youtube.com/watch?v=NHyl04-ytH8>>. Acesso em: 14 dez. 2011.

<sup>66</sup>Disponível em: <<http://www.youtube.com/watch?v=pdFJTbaFcZ0>>. Acesso em: 14 dez. 2011.

<sup>67</sup>Disponível em <<http://www.youtube.com/watch?v=2BrLEuzVCVQ>>. Acesso em: 14 dez. 2011.

<sup>68</sup>Disponível em: <<http://www.youtube.com/watch?v=aiXR9ggRdFI>>. Acesso em: 14 dez. 2011.

<sup>69</sup>Disponível em: <<http://www.youtube.com/watch?v=5rkoatnn9kM>>. Acesso em: 17 nov. 2011.

<sup>70</sup>Disponível em: <<http://www.youtube.com/watch?v=UUDwGFYS40I>>. Acesso em: 17 nov. 2011.

vibratos) e as palhetadas e dedilhados rápidos o suficiente para mal reconhecermos as notas que estão soando.

BluesMK, então, corta o áudio desta performance e a dubla, gravando um arquivo de áudio “sincronizado” enquanto assiste o vídeo original mudo, mimetizando Zappa. Nota-se na técnica de BluesMK a reprodução precária das técnicas de Zappa, mas indo além e forçando a saída do campo tonal para muito além do que Zappa foi originalmente, tocando aleatoriamente qualquer nota no braço da guitarra. Para intensificar o tom de comicidade, em meio à performance dublada, BluesMK inseriu, junto à sua versão sonora do solo, trechos da trilha sonora de abertura do filme *Guerras nas Estrelas* e um riff de uma canção da banda *Iron Maiden*. Mesmo assim, alguns usuários aparentemente compreenderam o *spoof* de BluesMK como uma performance “real” de Zappa e ainda por cima defendem o ilustre guitarrista, nos comentários do vídeo, dizendo que parte de seu trabalho era tentar não ser levado a sério. Não podemos saber com certeza se estes usuários estavam ou não cientes de que se tratava de uma dublagem, mas os cinco comentários abaixo (arranjados em ordem cronológica inversa), recortados dentre um total de 1.817 em 16 de novembro de 2011, tem muito a dizer sobre os *spoofs* de *shreds*:

Esse é mesmo Zappa tocando... Vocês tem que se dar conta de que Frank estava numa onda muito diferente da maioria de nós e ele tem um senso de humor muito seco. Basicamente seu humor está te desafiando mentalmente a ver isto como humor. E como ele disse muitas vezes “não me leve muito a sério ou você arruinará a performance”. (toobmaniac, 10/2011, tradução do autor)<sup>71</sup>

Não acredito que algumas pessoas pensam que isso é real... (AerosmithPower, 9/2011, tradução do autor)<sup>72</sup>

Um dos piores solos de guitarra tocado por um músico profissional que eu já ouvi... reprovado!!! (fatstrat6969, 7/2011, tradução do autor)<sup>73</sup>

Para todos aqueles que não sacaram, eles colocam um vídeo do Frank Zappa tocando e eles tocam horrivelmente por cima. (Vomuseind, 7/2011, tradução do autor)<sup>74</sup>

Interessante – Apesar dele estar fazendo isso como paródia, seu inconfundível uso de escalas e arpejos ainda aparecem nesse solo. (YesOuiSi, 2008, tradução do autor)<sup>75</sup>

<sup>71</sup>“That is very much indeed Zappa playing that. You have to realize that Frank was on a totally different wavelength than most of us are and he's got a very very dry sense of humor. Basically his humor is mentally challenging you to see it as humor. And as he said many times 'don't take me too seriously or you will ruin the performance’”.

<sup>72</sup>“Can't believe some guys think this is real..”

<sup>73</sup>“some of the worst guitar playing by a professional musician I have ever heard....dislike!!!!”

<sup>74</sup>“For all of those who did not get it, they put a video of Frank Zappa playing and they played horribly over it”.

<sup>75</sup>“Interesting --- Even though he's doing it as a parody, his unique one-of-a-kind use of scales and arpeggios still comes out a lot in that solo.”

Talvez o que coloca o espectador em dúvida é menos a desterritorialização da figura e da memória de Zappa, e mais a reterritorialização de algumas de suas técnicas. Comparando sonoramente, BluesMK foi muito prudente em imitar algumas técnicas de Zappa, e é por isso que o vídeo é potencialmente verossímil. O que ele faz é procurar elementos estruturais da virtuosidade de Frank Zappa para simular um território. Não é porque Zappa criou suas próprias idiossincrasias e assinaturas sonoras que alguém não possa tentar falsificá-las ou até instaurar diferenças nela. “Um território está sempre em vias de desterritorialização, ao menos potencial, em vias de passar a outros agenciamentos, mesmo que o outro agenciamento opere uma reterritorialização [...]” (DELEUZE; GUATTARI, 1997, p. 137). Há uma destituição de sentido por um lado, quando BluesMK rompe com a articulação entre a performance de Zappa e o áudio original, mas outros emergem pelo fato de o autor do vídeo produzir novas síncries entre o visual e o sonoro e simular a técnica de Frank Zappa. A seguinte citação de Chion pode explicar as diversas reações dos usuários nos comentários do vídeo:

Um mesmo som pode, segundo o contexto dramático e visual, contar coisas muito diferentes, uma vez que [...] mais do que o realismo acústico, é sobretudo o critério de sincronismo, e secundariamente de verosimilhança (verosimilhança que tem a ver não com o realismo, mas com a convenção), que o levará a colar um som a um acontecimento ou a um fenômeno. (2008, p. 24-25)

Outro aspecto interessante de salientar é que a maior parte do tempo do vídeo apresenta, no panorama visual, apenas Frank Zappa sozinho, conforme vemos no *frame* 31, mesmo que haja uma banda inteira de apoio, como podemos ouvir fora-de-quadro no vídeo original. BluesMK, no entanto, opta por gravar apenas uma dublagem da guitarra de Zappa. E pode ser essa uma das respostas de porque sua trilha visual fica tão aberta a possíveis reinterpretações sonoras, tanto que a versão de BluesMK cortou o som de todos os músicos de apoio, sem causar estranhamento.



*Frame 31*

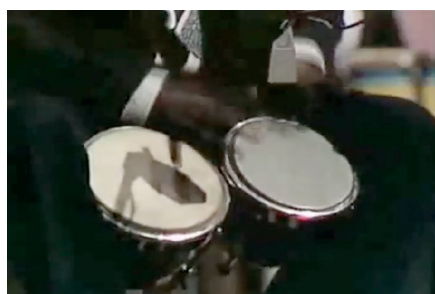
Os *spoofs* de *shreds* persistiram, e práticas derivadas foram surgindo. Uma performance de Paco de Lucia que sofreu *spoof* também nos chama atenção, porque assim como essa de Zappa foi completada com sons apenas que se referiam ao que aparece no panorama visual, no *Shred* de Paco de Lucia<sup>76</sup>, StSander tomou o mesmo caminho: embora, no vídeo original<sup>77</sup>, podemos ouvir os músicos tocando junto, o autor deste *spoof* preferiu dublar sonoramente apenas os instrumentos dentro do quadro. Assim, quando o enquadramento é similar ao do *frame* 32, apenas ouvimos bongôs e violão; no 33, um som oco, simbolizando como se Paco de Lucia tivesse golpeado o corpo do violão (o que não ocorre na trilha sonora do vídeo original); no 34, apenas bongôs, e enfim, no 35, a banda completa.



Frame 32



Frame 33



Frame 34



Frame 35

Vale lembrar que novamente a dublagem apresenta sons dos mesmos instrumentos, mas sendo tocados aleatoriamente, respeitando apenas os movimentos gestuais dos músicos na trilha visual. A comicidade está em reproduzir sonoramente apenas os instrumentos que aparecem visualmente, o que resulta numa confusão sonora.

Mas os sons nos Shreds não são só dos instrumentos musicais. Em alguns momentos, ouvem-se ruídos que simulam uma plateia apática. Algumas palmas e risadas, que podemos imaginar vir de no máximo meia dúzia de pessoas, completam a paisagem sonora. Nos vídeos originais, não se ouve a plateia e em alguns casos sequer há público. Aliás, essas

<sup>76</sup>Disponível em: <<http://www.youtube.com/watch?v=89zM9pZzt0U>>. Acesso em: 20 nov. 2011.

<sup>77</sup>Disponível em: <<http://www.youtube.com/watch?v=2oyhld64-s>>. Acesso em: 22 nov. 2011.

performances antigas mostram o público em pouquíssimos momentos, o que ajuda os *spoofs* a simular através da trilha sonora que há poucas pessoas assistindo, o que intensifica a comicidade do vídeo.

Isso nos ajuda a perceber como, historicamente, o audiovisual sempre tratou mais de exibir a performance dos músicos do que atribuir uma musicalidade estruturante às imagens visuais de concertos ao vivo. Isso não é nenhuma novidade, mas interessa perceber como a noção de sínkrise está muito mais relacionada com a tentativa de fazer um evento audiovisual parecer mais realista do que imprimir uma ritmicidade, uma harmonização ou um efeito de contraponto, sensações que seriam, naturalmente, derivadas da virtualidade da Música. E quando a televisão e o cinema tentam fazer isso, fracassam porque geralmente fazem tal apropriação nos seus termos, o que novamente demonstra como os ideais de Eisenstein não funcionam da mesma maneira como ocorrem na Música. A questão é que o audiovisual tomou de assalto vários termos da Música e os fez servir aos seus propósitos: leitmotiv, ritmo, harmonia, contraponto, marcação, tensão, clímax... Em compensação, o que estes vídeos para *web* parecem fazer é permitir que a Música exerça suas normas da forma “mais musical” possível. E é desse modo que estes vídeos de *Spoofs* de *Shreds* fazem emergir imagem-música: na atualização de elementos estruturais da Música no audiovisual, no acoplamento estrutural que este promove, mas *persistindo a sua musicalidade*.

Por fim, queria chegar a um vídeo de *shred* diferente de todos os outros, que se chama-se *John Williams Shreds*<sup>78</sup>. John Williams é um famoso compositor, autor dos temas de *E.T.*, *Parque dos Dinossauros*, *Guerra nas Estrelas* e *Indiana Jones*. No vídeo em questão, ele rege uma orquestra a tocar a poesia sinfônica *Assim Falou Zarathustra*, de Richard Strauss. Assim como nos casos precedentes, o autor do vídeo, o usuário Dayofthedave, cortou o áudio original e reproduziu sonoramente toda uma orquestra executando muito mal a peça. Claro que não se trata de um *shred*, pois não há guitarras, tampouco solos virtuosísticos, mas o termo virou sinônimo de dublar mal qualquer performance musical e publicá-la no *YouTube*. Na trilha sonora, Dayofthedave não toca aleatoriamente notas como nos casos de *spoofs* de Frank Zappa e Paco de Lucia, mas toca muito mal a mesma *Assim Falou Zarathustra*. Na maior parte do tempo, ouvem-se notas erradas, mas próximas da correta, enquanto vemos instrumentistas extremamente concentrados. A articulação entre o visual e o sonoro já é por si interessante, pois promove sínkrises similares aos vídeos mencionados anteriormente, mas, no entanto, me parece que algo ainda mais potente acontece aqui para o espectador, e que vai além da articulação entre visual e sonoro que vimos até aqui nos exemplos anteriores:

<sup>78</sup>Disponível em: <[http://www.youtube.com/watch?v=UuJrg\\_U8SNY](http://www.youtube.com/watch?v=UuJrg_U8SNY)>. Acesso em: 1 dez. 2011.



acontece que o tempo todo percebe-se, por mais mal tocado que esteja sendo, que se trata da *Zarathustra* de Strauss. E a irritabilidade que esta versão provoca se dá porque ela age diretamente na memória que temos da peça sendo executada apropriadamente. E, por algum mecanismo cognitivo, unimos a esta performance as imagens que temos da original.

A “harmonização” de ambos os fluxos – a da percepção que temos do vídeo ao ser executado e a das imagens que temos da execução adequada da peça – apresenta intervalos de sétimas e de segundas, que são as variações de desafinação que a versão de Dayofthedave faz em vários momentos. Quero dizer com isso que, se em algum momento a nota a ser tocada deveria ser um Dó, Dayofthedave foge dela, tendendo mais para o Si ou para o Ré, respectivamente a sétima e a segunda nota da escala em relação ao Dó. Dentro do sistema tonal, a utilização destas notas é feita para produzir uma sensação de proximidade – não é à toa que são chamadas de *sensíveis*. Quando as bandas de rock partem para as baladas, invariavelmente adicionam estas notas nos acordes. O jazz, por sua vez, tem essas notas inseridas em seu vocabulário o tempo todo. Se for para produzir essa sensação de aproximação, de sensualidade, esse efeito funcionará melhor diacronicamente – é mais potente, portanto, melodicamente: um Si que se sustente por mais tempo ao final de um compasso, atingindo o Dó ao final, produzirá tal sensação. Já sincronicamente, o efeito provocado é de tensão, pois a razão de vibração entre um Si e um Dó, ou qualquer intervalo de semitom na música, é de 15/16, de acordo com Wisnik (1999), mais complexo do que o de Dó e Fá, por exemplo, que é de 3/4<sup>79</sup>.

Ao ouvir a *Zarathustra* sendo mal executada, automaticamente acionamos na memória imagens-lembrança de como ela deveria ser executada corretamente, e a justaposição que fazemos na percepção da imagem-lembrança com as imagens sonoras que o vídeo nos promove exerce essa sincronia de notas sensíveis entre si, produzindo uma tensão apenas articulando imagens umas com as outras na nossa percepção. É claro que não ocorre uma realização de ondas sonoras em choque, mas tal evento torna-se potente na relação que nossa percepção faz entre as imagens – e tal pode ser comprovado pela aflição que sentimos ao assistir ao vídeo<sup>80</sup>. Aqui Bergson é muito relevante para a compreensão desse fenômeno, uma vez que reconhecemos uma tensão entre duas durações: das imagens sonoras recém processadas pelo vídeo e das imagens-lembrança da *Zarathustra* de Strauss executada adequadamente. Bergson explica como buscamos lembranças para agir no presente:

Temos consciência de um ato *sui generis* pelo qual nos afastamos do presente para

<sup>79</sup>Essas razões indicam quantos pulsos sonoros cada nota deve produzir para que ambas entrem em consonância.

<sup>80</sup>Algumas pessoas para as quais o presente não suportaram assisti-lo por inteiro.

nos recolocarmos, primeiro no passado em geral e depois numa certa região do passado [...]. Mas nossa lembrança continua em estado virtual; dispomo-nos assim apenas a recebê-la adotando a atitude apropriada. Pouco a pouco, ela aparece como uma névoa que se condensasse; de virtual, passa ao estado atual; e à medida que seus contornos vão se desenhando e sua superfície vai ganhando cor, tende a imitar a percepção. (2006c, p. 48-49)

E justamente por se atualizar, a virtualidade dessa lembrança da *Zarathustra* se torna mais complexa, além de ser irritada pela versão para *YouTube* de Dayofthedave. Essa lembrança faz parte do virtual da Música. E a música que estamos ouvindo ao ser executada em *John Williams Shreds* é potencialmente mais agonizante por estarmos também vendo uma orquestra aparentemente séria, dedicada, sendo regida por um compositor tão renomado, executá-la de forma tão descabida. Uma vez que a lembrança da Zarathustra bem tocada é puxada do fundo da memória pela percepção, “A lembrança aparece duplicando a cada instante a percepção, nascendo com ela, desenvolvendo-se ao mesmo tempo em que ela e sobrevivendo a ela, precisamente porque é de outra natureza” (2006c, p. 52). E no encontro dessas durações emerge uma imagem-música, e que é apenas potencial, pois nada mais é do que a relação entre duas outras imagens, uma gerada pela Música, a outra pelo audiovisual. Entre ambas surge essa tensão que irrita não só os ouvidos, mas a memória e toda a virtualidade da Música. A *Zarathustra* de Strauss está na nossa memória, está mais do que territorializada pelos inúmeros comerciais de TV, pelos filmes como *2001: uma odisseia no espaço* (Stanley Kubrick, 1968) e por filmes de pastiche que usaram a introdução da peça. É por isso que a versão desafinada nos agoniza tanto. Bergson nos dá mais pistas sobre o assunto.

A lembrança de uma sensação é coisa capaz de *sugerir* essa sensação, ou seja, de fazê-la renascer, fraca primeiro, mais forte em seguida, cada vez mais forte à medida que a atenção se fixa mais nela. Mas a lembrança é diferente do estado que sugere e é precisamente porque a sentimos por trás da sensação sugerida [...]. [...] a lembrança pura de uma sensação ou de uma percepção não é, em nenhum grau, a sensação ou a percepção elas mesmas. (2006c, p. 51)

Então, porque somos capazes de diferenciar percepção de lembrança é que é possível a emergência dessa imagem-música entre as duas durações. A imagem da simultaneidade das duas performances, que ocorre em um fluxo único e exclusivo da percepção do espectador, que é forçado a imaginar um encontro musical que nunca se manifestou.

Os *spoofs* de Shreds apresentam, portanto, outros modos de expressar potencialidades musicais através de audiovisualidades: ao contrário do que ocorre com os *VideoSongs*, em que o processo de edição deixa evidente que a canção foi pré-concebida, os *spoofs* de *shreds*, especialmente suas trilhas sonoras, são uma expressão que só pode derivar da leitura do editor

do vídeo de potencialidades musicais em devir na trilha visual de vídeos que apresentam músicos em performance.

E se a performance visual ou a música executada for uma que consta na memória coletiva, como a *Zarathustra* de Strauss, ela ainda é capaz de desterritorializar a imagem-lembrança que temos de uma performance adequada da peça, fazendo com que essa imagem precise ser resgatada pela percepção para poder agir no presente, num esforço para compreender as imagens visuais e sonoras provenientes do fluxo audiovisual.

#### 5.4 AUTO-TUNNING

As considerações sobre vídeos que trabalham com a ideia *sampling* audiovisual, mencionados em 4.2., aplicam-se também ao tipo de vídeo que será analisado neste trecho.

Segundo o *site Know Your Meme*, a empresa de software Antares Audio Technologies desenvolveu em 1997 o plugin Auto-Tune, que processa automaticamente a afinação de qualquer onda sonora<sup>81</sup>. Inicialmente utilizado em estúdios profissionais para corrigir os deslizes dos cantores modernos, a cantora Cher o utilizaria um ano depois para seu single *Believe*, em que ela usou o software para forçar este efeito, evidenciando-o como uma “sujeira” sonora. Seu uso na música popular hoje é bastante recorrente. O *rapper* T-Pain chegou a gravar um disco inteiro usando o software para melodizar sua voz. Ke\$ha e Britney Spears também usam o efeito com certa frequência. Com a popularização do software, usuários do *YouTube* passaram a aplicar o efeito para melodizar todo tipo de áudio (dentre alguns, miados de gato, choros de bebês, personagens de videogame), e publicá-los em vídeo no *site*. Como se não bastasse, na trilha visual as imagens acompanham o desenrolar dos eventos sonoros melodizados. Um dos primeiros vídeos de *auto-tunnig* chama-se *Winston Churchill backed by band from the future*<sup>82</sup>, publicado em fevereiro de 2009. Vemos o autor Michael Gregory (nome de usuário *schmoyoho*) rufando um tambor que o faz “viajar no tempo”, ao que, em seguida, no próximo plano, vemos uma imagem de arquivo de Churchill falando aos aliados em 1941, com Gregory inserido em cromaqui como se estivesse também sentado à mesa. Com a voz passada por Auto-Tune, Churchill agora participa de uma canção pop com instrumentos musicais sintetizados, o que demonstra mais um poder das remixagens:

<sup>81</sup>Fonte: <http://knowyourmeme.com/memes/auto-tune>. Acesso em: 26 nov. 2011.

<sup>82</sup>Disponível em: <<http://www.youtube.com/watch?v=IW6jW9y59JY>>. Acesso em: 26 nov. 2011.

[...] pode-se levantar a hipótese de que, assim como os Disk Jockeys (DJs) têm recriado antigas canções, conferindo-lhes forma digital, para o que transformam as gravações originais em matérias para nova formalização, também parece pertinente imaginar a possibilidade de se inventarem audiovisuais digitais a partir da matéria audiovisual disponível, trazendo de novo à cena viva, como matéria, o que parecia estar definitivamente calado. (SILVA, 2006, p. 47-48)<sup>83</sup>

Michael Gregory também tem parceria com outros músicos/videastas, que formam *The Gregory Brothers*. Juntos, produzem desde 2009 a série para *YouTube Auto-Tune the News*, em que extraem as falas de âncoras de TV e outras “cabeças falantes” e os *auto-tunam* para que façam parte de uma música por eles composta, elaborando discursos sem-sentido pelo cruzamento de matérias jornalísticas que não possuem relação alguma umas com as outras.

Assim como nos *samplers* audiovisuais, os compositores acabam por elaborar introduções, repetições de falas para criar refrões, e tudo o mais que a Música é capaz de forçar. A imposição do Sistema Musical em *Auto-Tune the News* é tal que em todos os episódios as falas extraídas dos noticiários, mesmo “auto-tunadas”, precisam de acompanhamento instrumental. No *Frame 36*, nota-se que os produtores se inseriram na imagem via cromaqui, tocando bateria e guitarra<sup>84</sup>. Essa limitação criativa é justamente o que potencializa a série: a Música, ao observar o audiovisual, o força a fazer acoplamento estrutural, não só na trilha sonora, mas também na visual.



*Frame 36*

O processo de edição não é simples: de cada plano audiovisual retirado do noticiário, o áudio das falas precisa ser separado e tratado em software de produção musical, onde passam pelo tratamento de melodização (*auto-tune*) e sincronização com a trilha instrumental. Em seguida, esse áudio é importado para um software de edição não-linear de vídeo, onde cada

<sup>83</sup>Tal constatação de Silva também se aplica aos *Mashups* de Chaves abordados em 5.1.

<sup>84</sup>Disponível em: <<http://www.youtube.com/watch?v=qizNQKzatXA>>. Acesso em 13 jan. 2012.

plano visual referente a cada fala é então ressincronizado. Todas essas operações sem dúvida marcam um novo estágio do processo composicional musical, que se funde com outro processo de composição, audiovisual, numa intensidade de sincronias nunca antes atingida, pois, ao contrário da trilha sonora no cinema e da do videoclipe, não é possível compor a canção antes do audiovisual. Cada interferência na estrutura da canção na trilha sonora influenciará no resultado das imagens visuais.

Como cada fala é picotada para ser sincronizada à canção, e sua respectiva imagem precisa estar dela acompanhada, o que temos é a emergência de uma música bem desenhada e fluida, acompanhada por imagens desconexas umas das outras (tal como ocorre também nos *VideoSongs*), não por uma opção estética como ocorria nos anos 80 e 90 nos videoclipes, mas por consequência de uma montagem que quer dar a ver uma música que estrutura as audiovisuais. O visual, as imagens picotadas, ao contrário da canção resultante na trilha sonora, tende para a periferia da Música, onde ocorrem os embates para aceitar ou refutar elementos irritantes ao Sistema. São imagens potencialmente dotadas de musicalidade.

É uma virada de jogo, via uma observação de segunda ordem: os compositores (aqueles do sistema “processo composicional” que vinham sendo exemplificados no capítulo 4.2) observam o audiovisual e procuram *pontos brilhantes*, momentos do telejornal que seriam mais interessantes de serem *auto-tunados*, enquanto que a Música observa a observação dos compositores, exigindo que suas estruturas sejam aplicadas no material.

O vídeo mais assistido e repercutido dos Gregory Brothers foi publicado em 31 de julho de 2010: o vídeo *Bed intruder song!!!*<sup>85</sup>. É uma remixagem do vídeo *Antoine Dodson warns a PERP on LIVE TV!*<sup>86</sup>, que se trata de um recorte de uma reportagem sobre um pervertido que tentou invadir uma casa e estuprar uma mulher. Nessa reportagem, o irmão da vítima, Antoine Dodson, é entrevistado e faz ameaças ao invasor diante das câmeras. Seus trejeitos, modo de vestir e performance lembram muito os cantores de *hip hop*. A partir de frases de efeito de Dodson, os irmãos Gregory produziram uma canção melodizando as frases do rapaz com o Auto-Tune. Interessante notar que nesse caso em especial, ao contrário das falas monocórdias de Winston Churchill e dos âncoras de noticiários usados na série *Auto-Tune The News*, Dodson desenvolve curvas melódicas carregadas de ritmicidade. O que os irmãos Gregory tiveram de fazer foi encaixar os ataques das sílabas com um ritmo regular, controlado por metrônomo digital, e afinar a voz dele com o Auto-Tune. O curioso é que a aplicação do *software* nesse caso lembra muito mais os primeiros usos, pois a harmonização

<sup>85</sup>Disponível em: <<http://www.youtube.com/watch?v=hMtZfW2z9dw>>. Acesso em: 27 nov. 2011.

<sup>86</sup>Disponível em: <<http://www.youtube.com/watch?v=EzNhaLUT520>>. Acesso em: 27 nov. 2011.

ficou tão próxima da melodia sugerida pela fala de Dodson que o *auto-tunning* ficou discreto. Assim como em *MC Jeremias e Funk da menina pastora*, em que os autores do remix perceberam um devir funk carioca na fala de Jeremias e da pastora mirim, e também como os personagens do seriado *Chaves* têm um devir heavy-metal nas suas performances visuais, Antoine Dodson está carregado de uma potência *hip hop*.

Mas *Bed intruder song!!!* passa por procedimentos mais complexos que os anteriormente analisados. Para encaixar as sílabas no tempo, em alguns momentos os planos audiovisuais passam em *slow* ou *fast motion*, dependendo de quanto elas precisam durar. Se este efeito fosse produzido em fita cassete, o *slow motion* faria a voz ficar mais grave, enquanto que o *fast*, mais agudo. No entanto, como o Auto-Tune é aplicado (somado a outros *plugins* que mantêm a altura melódica mesmo que se modifique a duração da trilha sonora), torna-se possível aumentar ou diminuir o tempo de duração dos *samplers* audiovisuais sem interferir na curva melódica. Nesse caso, o visual e o sonoro andam muito mais juntos do que nos *VideoSongs*, nos quais é possível que os músicos/videastas exibam uma imagem visual que não foi registrada no mesmo momento que a sonora e vice-versa, embora afirmem na descrição que o visual e o sonoro sempre provém de uma mesma captação. Em *Bed intruder song!!!*, temos acesso ao material original (a reportagem sobre a tentativa de estupro), e podemos constatar que todas as imagens (visuais e sonoras, é preciso lembrar) usadas na remixagem derivam de outro audiovisual.

Penso que os vídeos de Auto-Tune, especialmente os que se alinham ao método dos irmãos Gregory, expressam muito bem o que quero entender aqui como imagem-música, porque eles não tratam apenas de musicalizar qualquer evento sonoro, mas de musicalizar um audiovisual, sua trilha visual e sonora simultânea e indistintamente. O material original já é potencialmente musical, seja na trilha sonora, porque ouve-se alguém como Antoine Dodson praticamente cantando enquanto fala, seja na trilha visual, porque um mesmo personagem como esse incorpora elementos da virtualidade da Música como postura, moda, movimentos e performances.

É curioso também perceber que o resultado das montagens tende para as estruturas da canção popular. Se as coisas soam melhor afinadas, é porque elas tendem para o epicentro estrutural da Música, para uma zona de conforto, o “museu” da Música. O que podemos cogitar é que a montagem e a composição em si já são carregadas demais de relações, de imagens, de interferências, e que são mais fáceis de serem reconhecidas se a canção resultante for fácil de ouvir. Se o resultado fosse uma peça dodecafônica talvez fosse informação demais para nossa percepção, e acabaríamos inevitavelmente tentando prestar atenção ora no visual,

ora no sonoro. Ao contrário, os vídeos de *auto-tunning* parecem ser compostos de tal forma que os processos de montagem se evidenciem aos olhos do espectador. Então ficam muito longe de serem da ordem da imagem-movimento, pois, embora o material original derive de telejornais que operam com essa lógica, o fracionamento, o *sampling* de trechos deste vídeo o converte em situações óticas e sonoras puras, nos termos da imagem-tempo. Mas não para aí: esses *samplers* dão mais um passo à frente para se transformarem em situações musicais. Tal situação, nos termos da imagem-música, ainda é carregada das outras ordens. Não se tenta simular como se Antoine Dodson estivesse cantando, não se tenta produzir uma verossimilhança. Pelo contrário, a montagem *se mostra*, nos moldes da imagem-tempo, porque é evidente a falsificação, a evidenciação dos processos de produção é que dão o toque de comicidade a essas imagens, do mesmo modo como falei sobre o *MC Jeremias* no capítulo 2.1. E só podemos perceber que há *sampling* visualmente. Enquanto é muito fácil percebermos quando há uma troca de planos na trilha visual, isto se torna muito difícil na trilha sonora, pois, de acordo com Chion,

a montagem de sons no cinema não criou uma unidade específica. Não percebemos unidades de montagem-som, e as colagens sonoras não nos saltam ao ouvido nem nos permitem delimitar blocos identificáveis. Isso não é, de resto, exclusivo do cinema: montam-se sons desde que a técnica o permite (quer dizer, desde cerca dos anos 1930), tanto no rádio e em disco como na música em fita. Ora, em nenhum destes casos, haja ou não imagem, a noção de plano sonoro como unidade de montagem conseguiu emergir como unidade neutra e reconhecida por todos (2008, p. 39)

O *auto-tunning* é oposto à ideia da montagem separada da trilha visual e sonora, como Chion apresenta. Não se trata de disfarçar a montagem sonora, mas de evidenciá-la *através* da montagem visual, uma vez que se percebe que imagem visual e sonora foram capturadas simultaneamente. A imagem-música emerge menos dessa relação entre o visual e o sonoro, e mais entre uma música potencial e a montagem audiovisual.

É nesse tipo de vídeo que fica mais clara a manifestação da Música, e também é o que, dentre os quatro tipos encontrados, mais se distancia da lógica do videoclipe. A diferença provocada na Música pelos vídeos de *auto-tunning* se dá menos na canção manifestada sonoramente, e muito mais através dos processos composicionais, que utilizam como matéria prima apenas outros vídeos que já haviam sido publicados no *YouTube*, que não são musicais, e que são remixados. É diferente da proposta dos *spoofs* de *shreds*, nos quais a dublagem é pensada e composta pelo compositor do vídeo, que tem certa liberdade de que notas irá escolher para tocar; diferente também dos *VideoSongs* da banda Pomplamoose, que, assim como *Amateur*, utilizam como matéria prima vídeos capturados pelos próprios compositores –

sabe-se que a música já foi preconcebida em ambos os casos antes de se iniciar a edição do vídeo, talvez até tenha sido ensaiada ao vivo; o que dizer dos *mashups* audiovisuais então, que também partem de músicas pré-concebidas, quase como nos moldes dos videoclipes?!

No caso dos vídeos de *auto-tunning*, o resultado final, a canção, deriva do que os vídeos pré-publicados no *YouTube* sugerem melodicamente. E podemos ainda assistir aos vídeos originais para reconhecer a música em potência que foi trabalhada pelos compositores. Eles nos ensinam a *imaginar* Música no audiovisual, a reconhecer as potencialidades Musicais de telejornais, documentários, *reality shows* e, quiçá, videoclipes.

Eles sugerem ao espectador que faça outra leitura de onde ele deve reconhecer música: não mais apenas como uma manifestação sonora acabada, mas como uma virtualidade que problematiza o audiovisual.

## 5.5 4'33"

O último tipo de vídeo é um bem específico, e se após a leitura das análises dos *mashups* e *sampling* audiovisual, dos *spoofs de shreds* e dos vídeos de *auto-tunning* pôde-se constatar vários pontos em comum entre estes, o que será analisado a seguir diferencia-se bastante dos demais. Este subcapítulo é quase um apêndice, mas precisa ser incorporado aqui porque ele retoma várias questões que foram sugeridas e ensaiadas nos outros capítulos. Agora, os amarro a um tipo de vídeo musical para *web* que, justamente por fazer emergir imagens-música, não pode ser encarado como apêndice. Esta última análise trata de vídeos cujos títulos sugerem ser execuções audiovisuais da peça 4'33" de John Cage. Tanto a peça quanto seu compositor apareceram ao longo da dissertação problematizando o objeto de pesquisa. Daí a importância de seu retorno, agora, como objeto de análise<sup>87</sup>. Começemos por um depoimento do próprio compositor sobre a *première*, executada em 1952 por David Tudor:

As pessoas começaram a cochichar umas com as outras, e algumas foram embora. Eles não deram risada — eles somente se irritaram quando se deram conta de que nada iria acontecer, e eles não se esqueceram disso 30 anos depois: eles ainda estão irritados. (tradução do autor)<sup>88</sup>

<sup>87</sup>Agradeço aqui a Gabriel Saikoski pelos longos embates e debates que tivemos em 2009 sobre esse tipo de vídeo.

<sup>88</sup>Cage conversation with Michael John White (1982), in Kostelanetz 1988, 66, in: Solomon, Larry J.: The Sounds of Silence, in: <http://www.azstarnet.com/~solo/4min33se.htm>. No original: "People began whispering to one another, and some people began to walk out. They didn't laugh – they were just irritated when they realized nothing was going to happen, and they haven't forgotten it 30 years later: they're still angry."



Já faz 60 anos da *première* de *4'33''* e as pessoas ainda estão irritadas. Provavelmente, assistir à execução desta peça sem saber previamente do que se trata deva ser realmente desagradável. No primeiro contato que tive com ela, já sabia do que se tratava, então pude fruir com mais serenidade. Ou será que não o fiz corretamente? Talvez Cage me corrigisse, dizendo que não se trata de ouvir corretamente ou não, e sim que a peça procura romper com a ideia de sujeito e objeto. Não quer dizer que, pelo fato de os músicos não executarem som algum com seus instrumentos, a peça seja dotada de silêncio, mas que tudo ao redor (a plateia, o ar condicionado, o som do trânsito) *é* a música. Pode-se dizer, portanto, que *4'33''* é composta justamente por aquilo que ela não é.

Mas talvez Cage respondesse minha pergunta dizendo “por favor, quer repetir a pergunta?” (ECO, 2007, p. 212), e repetir essa resposta a cada vez que eu repetisse a pergunta; ou ainda ele poderia responder com uma frase como “a nuvem tapa o sol”, ou ainda me dando um bofetão na cara. Isso porque o compositor tinha uma relação muito forte com o zen-budismo, cuja proposta trata de fazer com que deixemos de habitar os conceitos que criamos do mundo para voltar a habitar o próprio mundo.

Várias características da peça têm relação com a Teoria dos Sistemas de Nicklas Luhmann. Se o espectador dá um espirro durante *4'33''*, ele está simultaneamente observando e sendo observado pela peça. Se, para os conceitos da música tradicional, um espirro seria um ruído externo à peça, agora ele faz parte. Ela aceita todo tipo de irritação do meio, porque ela também é meio, ou melhor, ela *é o meio*. Ouvei *4'33''* com o entendimento de que a peça provocaria uma dissolução de sujeito e objeto, colocando-nos não só como espectadores, mas também como executores, alternando entre esses papéis. Não me lembro de ter ouvido os sons como tais: se houvesse um ar condicionado ligado, ao perceber o som, pensava “eis o som do ar condicionado, que, antes de executar a música, não me dei conta”. Não lembro tampouco de ter me sentido executor, pois ou tentava me manter silencioso, ou provocava sons propositadamente. Mas a peça trata dessa indiscernibilidade: se não conseguirmos nos conectar selvagememente com o mundo, ao menos ela serve para demonstrar como estamos o tempo todo tentando domesticá-lo. A peça irrita o espectador, mas esse, por sua vez, não consegue converter tal irritação em informação, pois *4'33''* não repete elementos, nem mesmo se ela for executada uma segunda vez: ela é composta da mais pura diferença possível de se extrair da Música. Ela não provoca territorializações, e é por isso que o espectador não pode extrair informações dela, apenas sensações (irritações, nos termos de Luhmann, 2009).

Introduzida a peça, tratarei agora dos vídeos. Fazendo a pesquisa exploratória no

*YouTube*, me deparei, surpreendentemente, com vídeos dos usuários executando 4'33". Mais do que uma homenagem, eles estendem para o mundo virtual a inquietação e a proposta de indissociabilidade de sujeito-objeto de John Cage. A grande maioria dos vídeos encontrados é dotada de certa comicidade. O que não chega a ser um problema, se lembrarmos que Cage não queria se levar a sério, tampouco queria produzir sentido através de suas obras, assim como um bom humor *non-sense* deve ser. Há tanto performances gravadas ao vivo quanto representações eletrônicas. Do primeiro tipo, há *How to Play 4'33" by John Cage*<sup>89</sup> (*Frame 35*), em que um músico jovem, diante de um teclado, apresenta um tutorial sarcástico de como tocar a peça. Nas suas palavras: “errando” a execução ao atacar as teclas acidentalmente em três momentos distintos. Há também uma performance executada dentro de um ônibus<sup>90</sup> em movimento (*Frame 36*). Em *A Rendition of John Cage's 4'33"*<sup>91</sup>, vemos dois músicos executando a peça em um pequeno palco, provavelmente de um bar, tentando não rir. Num outro vídeo, um grupo de adolescentes<sup>92</sup> executa a peça em uma sala de estar. Eu poderia apresentar exemplos intermináveis, mas encerro a amostragem com um último caso curioso: uma versão a capella<sup>93</sup> (*Frame 37*), em que um mesmo garoto se registra em seis momentos diferentes, dividindo a tela também em seis (prática comum em vídeos para *web*), como os *VideoSongs*. É possível encontrar no *YouTube* execuções com todo tipo de instrumentos<sup>94</sup>: piano, guitarra, baixo, bateria, copos de cristal, ukulele, flauta, órgão, marionete... Em todos os casos, na trilha visual comparecem os músicos que estão executando a peça em um ambiente, enquanto na trilha sonora estão registrados os sons que microfones captaram do mesmo ambiente. Alguns músicos explicam no início o que vão executar, outros apenas se registram diante da câmera, imóveis. Da mesma forma, muitos não seguem à risca as instruções da partitura original (como executada por David Tudor em 1952), como marcar com um relógio e dividir em três movimentos, mas, novamente, este era um desejo de Cage, que dizia que a peça podia durar qualquer tempo e ser tocada por qualquer instrumento.

---

<sup>89</sup>Disponível em: <<http://www.youtube.com/watch?v=8LJFJyvZA94>>. Acesso em: 1 dez. 2011.

<sup>90</sup>Disponível em: <<http://www.youtube.com/watch?v=WYQhXN1UFbU>>. Acesso em: 26 jul. 2011.

<sup>91</sup>Disponível em: <<http://www.youtube.com/watch?v=SkzU0RE6PzA>>. Acesso em : 9 out. 2011.

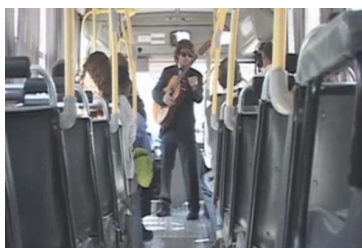
<sup>92</sup>Disponível em: <<http://www.youtube.com/watch?v=kyRpAIv9IUU#!>>. Acesso em: 9 out. 2011.

<sup>93</sup>Disponível em: <[http://www.youtube.com/watch?v=A4hsBB\\_g9vM](http://www.youtube.com/watch?v=A4hsBB_g9vM)>. Acesso em: 9 out. 2011.

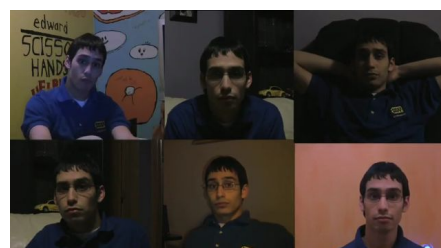
<sup>94</sup>O usuário Dick Whyte postou no YouTube um vídeo em que compila trechos de quatro segundos de duração de 68 outros vídeos, também publicados no YouTube, que exibem representações de 4'33" das mais diferentes maneiras. O vídeo está disponível em: <[http://www.youtube.com/watch?v=8tHE\\_52ae2s](http://www.youtube.com/watch?v=8tHE_52ae2s)>. Acesso em: 9 out. 2011.



Frame 37

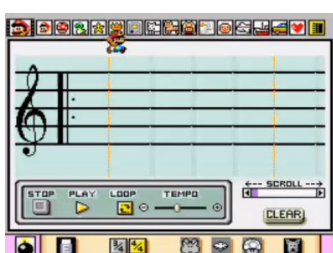


Frame 38



Frame 39

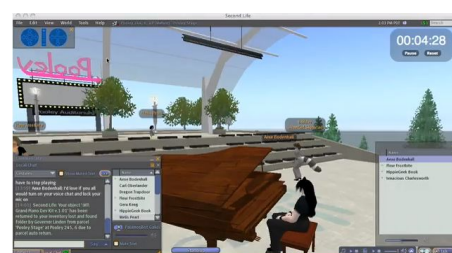
As representações eletrônicas são ainda mais curiosas: Em 1992, a Nintendo lançou o cartucho *Mario Paint* para o console Super Nintendo. Nele, tinha um compositor de partitura. Anos depois, uma versão emulada para computadores pessoais virou febre entre os internautas que montavam versões de várias canções e publicavam no *YouTube*. Eis que alguém cria, provavelmente, a primeira execução de 4'33" eletrônica orquestrada e regida por ninguém mais ninguém menos que Mario (o personagem ítalo-americano da Nintendo. *Frame 38*)<sup>95</sup>. Em *Moonfish playing John Cage's 4'33"*<sup>96</sup>, peixes de aquário, enquadrados na tela, estariam executando a peça (*Frame 39*). Encontrei também um vídeo de uma execução da peça na rede social *Second Life* (*Frame 40*) e outro em que um usuário simulou como seria a execução da peça no videogame *Guitar Hero*<sup>97</sup>, alternando dois planos: um de suas mãos repousando sobre o controle do jogo, em formato de guitarra, outro da tela do televisor, onde o característico braço de guitarra interminável, que corre verticalmente na tela durante as canções, mas sem nenhuma indicação de nota a ser tocada. Há até mesmo um vídeo cuja duração total é composta por tela preta e sem registro de áudio<sup>98</sup>.



Frame 40



Frame 41



Frame 42

A primeira questão que aparece é se estamos ou não diante de execuções *plenas* de 4'33". Como já disse, a peça pode ser executada por qualquer músico com qualquer instrumento e durar o tempo que ele quiser. Mas o músico precisa sinalizar o início e o fim da peça. Isso quer dizer que a peça necessita de uma performance, facilmente identificada da

<sup>95</sup>Disponível em: <[http://www.youtube.com/watch?v=TYFL\\_vxegEY](http://www.youtube.com/watch?v=TYFL_vxegEY)>. Acesso em: 9 out. 2011.

<sup>96</sup>Disponível em: <<http://www.youtube.com/watch?v=EPgFE8T8E2E>>. Acesso em: 9 out. 2011.

<sup>97</sup>Disponível em: <<http://www.youtube.com/watch?v=8XHYYmoQ7IJ8>>. Acesso em: 10 out. 2011.

<sup>98</sup>Disponível em: <[http://www.youtube.com/watch?v=6QUMYQEO\\_9g](http://www.youtube.com/watch?v=6QUMYQEO_9g)>. Acesso em: 9 out. 2011.

parte dos músicos inseridos nos vídeos, e não podemos negar que as representações eletrônicas também não deixam de ser performáticas, pois mesmo na ausência de um músico no panorama, o início e o fim do vídeo já são o recorte que pressupõe a performance. Reitero que não se trata de definir se são ou não execuções de 4'33". Mesmo que chegássemos a uma resposta, ela pouco diria respeito ao que estou querendo estudar. Importa, isso sim, evidenciar como a virtualidade da peça diferencia-se de si nestes vídeos.

Outra questão que surgiu a partir da observação empírica, que pode ajudar a circunscrever o objeto: É possível que haja uma reprodução técnica de 4'33"? Reproduzir uma obra musical, executar 4'33" ao vivo, por exemplo, é transformá-la num objeto para outros sujeitos poderem ouvi-la. Talvez seja possível dizer que, quando vemos um vídeo de um músico executando 4'33", nos tornamos um espécie de meta-sujeito. Somos sujeitos que veem um segundo sujeito (o câmera e/ou o editor do vídeo) ouvindo a performance (hetero-observação). O vídeo de Andres Torres, em que vemos um músico tocando violão em um ônibus, transforma o espectador em meta-espectador. Ao contrário da leitura tradicional de música, ela não é, unicamente, o som que sai dos instrumentos, tampouco o será dos alto-falantes de nossos computadores. Para sermos estritos, teremos que adicionar à trilha sonora do vídeo os sons que ouvimos no ambiente em que o vídeo estiver sendo exibido. Mas este não é um procedimento que o audiovisual é capaz de fazer, nem a análise aqui proposta. A dificuldade reside na impossibilidade de dizer *onde* está a peça. Talvez possa se dizer que, ao assisti-la, fazemos parte do vídeo também. Como vimos no capítulo 2.2, o cinema tratou de sistematicamente abstrair os sons que cercam uma exibição. Assistir uma performance de 4'33" em meio a um filme de ficção falharia, portanto, pois os espectadores recortariam os sons a serem ouvidos apenas aos que saíssem dos alto falantes. A peça, por sua vez, pressupõe uma conexão com o fluxo, com a duração bergsoniana (em oposição ao modo que normalmente assistimos audiovisuais, os recortando do ambiente ao redor como se fossem um fluxo à parte). Por isso, nos moldes em que assistimos a vídeos no *YouTube*, torna-se impossível reproduzir audiovisualmente a peça. Mas, por outro lado, ao invés de reprodução, podemos pensar em outro termo. Não me parece também possível compreender esses vídeos como performances plenas de 4'33"; sendo estrito, talvez eles nada tenham a ver com a peça. Mas o que importa aqui é perceber como o audiovisual reage a tal imposição da Música. Ao contrário de assistir uma execução dessa peça ao vivo, em que a paisagem sonora imediatamente se confunde com a execução, ao assistir um audiovisual nos esforçamos por ignorar os sons à volta, mesmo que este audiovisual não possua som algum. Além do mais, uma vez que eu os estou analisando, só posso falar do que ocorre na trilha sonora e na trilha

visual. Os sons que circundam a exibição destes vídeos o leitor terá que adicionar às análises aqui feitas, se desejar aproximar os vídeos a uma reprodução ideal de 4'33". Por isso, trataremos os vídeos como imagens técnicas da peça, como um conceito (melhor do que pensá-las como reproduções): como imagem-música. 4'33" está contida nesses vídeos, mas apenas potencialmente, por isso como imagem-música. Ela não se realiza plenamente. Não a reproduzem, não a representam, não a executam (vale dizer que essa impossibilidade de tradução plena é característica de qualquer tradução). O que esta dissertação pode sugerir é o seguinte: que esses vídeos *imaginam* uma execução audiovisual de 4'33", bem como Flusser sugere em *Elogio da Superficialidade*<sup>99</sup> (2008): sintetizando imagens, podemos hoje fazer *música imaginativa*. Através de imagens técnicas, vale dizer.

Se concordamos com isso, avanço para outra problemática. Ao contrário de uma temporada de concertos ao vivo, em que um músico executaria diariamente 4'33", cada execução sendo completamente diferente uma da outra, não podemos dizer o mesmo destes vídeos. Pois a trilha sonora é sempre a mesma quando executada. Mas se 4'33" é composta de diferenças, como um registro audiovisual dela pode dar conta dessa diferença? Estamos falando de uma obra aberta, “[...] uma categoria mais restrita de obras que, por sua capacidade de assumir diversas estruturas imprevistas, fisicamente irrealizadas, poderíamos definir como 'obras em movimento’” (ECO, 2007, p. 51). Ou talvez estejamos, numa maior aproximação, falando de uma obra indeterminada. E a indeterminação seria a chave zen-budista de Cage para atingir uma experiência do som antes de um significado do som. O serialismo de Stockhausen, em oposição, tenta ir ao encontro de uma música extremamente exata, matemática, chegando a extremos como o de solicitar, em partitura, que o pianista ataque as teclas do piano distribuindo sete dinâmicas diferentes pelos dedos da mão – uma exigência de ordem super-humana. Buscar tal exatidão matemática na música é uma tentativa de transformá-la em representação pura. E é isso que os vídeos permitem: a exibição repetida, idêntica, de uma peça cuja proposta é ser diferença. Cada vídeo, nesse *e somente nesse* ponto de vista, perde toda sua potência de diferença. Mas por outro lado, o compósito de vídeos potencializa a desterritorialização da peça. Isto é, eles propõem uma diferente leitura. À primeira vista, parece que ela é muito mais para ser vista do que ouvida no *YouTube*, pois ao recordar as descrições dos vídeos, o que os diferencia é “quem” aparece tocando ou “onde” estão tocando. Além disso, quase todos os que são aqui exemplificados se propõem cômicos, tentando criar constrangimentos ao espectador. Relembrando a entrevista que Cage concedeu

---

<sup>99</sup>*Elogio da Superficialidade* é o subtítulo de *O universo da imagens técnicas*, mas em seu original datilografado é o título, muito mais adequado para responder as questões que aqui apresento.

em 1991, no filme *Ecouté*, ele desejava que suas composições fossem o mais destituídas de sentido possível. Cage propõe que se esqueça do objetivo na escuta. Uma escuta sem memória, antes da representação, antes de qualquer dualismo. *Antes de qualquer relação entre imagens*. No entanto, todos os vídeos acabam agregando sentidos à peça, quando a proposta de Cage sempre foi oposta a isto. Ainda que o humor desses vídeos seja *non-sense*, a articulação dos eixos entre o sonoro (“silêncio”) e o visual (os músicos olhando fixadamente a câmera, às vezes se coçando, tossindo ou bocejando propositadamente) promove um sentido, que é o de *parecer* ser um vídeo sem sentido. E se rememorarmos a maioria dos outros casos, os *mashups* de Chaves com System Of A Down, os *auto-tunnings* e os *spoofs* de *shreds*, a proposta é a mesma: é destituir o sentido de vídeos de arquivo na sua remixagem, mas provocando outros sentidos ao rearranjá-los diacrônica e sincronicamente.

Se os vídeos fazem *4'33"* perder potência realizadora, por outro lado eles contribuem no entendimento de questões do audiovisual, ao ser irritado pela Música. A que me parece principal nesse caso é: como pode ser uma imagem do silêncio no audiovisual?

Certa vez, John Cage foi convidado a visitar uma câmara anecoica<sup>100</sup>. Os engenheiros lhe disseram ser das salas mais silenciosas do mundo. Ao adentrá-la, o compositor se surpreendeu com o efeito do ambiente. A absorção sonora das paredes acolchoadas era tamanha que ele sentia como se as palavras que proferia morriam logo que saíam de sua boca. No entanto, contando com uma audição de músico, foi capaz de ouvir dois sons na sala, e relatou esse fato ao engenheiro, que lhe perguntou: “como são os sons que você está ouvindo?”. Cage: “um é muito agudo, enquanto o outro é muito grave. O engenheiro lhe respondeu que o primeiro era o som do sistema nervoso de Cage em funcionamento, enquanto que o grave era o coração batendo. Depois desse evento, Cage pensou que não precisava mais se preocupar com o futuro da música, pois se deu conta de que não tem como o ser humano experienciar o silêncio. Adiciono aí que não se trata de dizer que *não existe silêncio*, mas que ele é uma abstração; talvez seja melhor dizer, o silêncio é um signo.

E *4'33"* trata exatamente disso. Pois bem, agora, quando assistimos aos vídeos da *web* que imaginam execuções da peça (como uma imagem técnica flusseriana), que “silêncio” eles produzem? Já que estou pressupondo que só posso falar da trilha visual e sonora, nelas podemos perceber a possibilidade de se produzir silêncio no audiovisual. Se nenhum pulso elétrico for enviado para as caixas acústicas, elas não são capazes de vibrar o ar, portanto, ao menos nesse recorte, é possível dizer que o audiovisual pode ficar em silêncio, e não será um

---

<sup>100</sup>O relato do próprio compositor sobre a experiência na câmara anecoica pode ser assistida na íntegra no filme *Four american composers – John Cage*, dirigido por Peter Greenaway, 1991.

silêncio abstrato. É um silêncio relativo, porque o ambiente e o próprio espectador estão produzindo sons. Mas se lembrarmos de Flusser, que diz que o espectador quer ser enganado, ele *abstrai* o resto do ambiente à volta e foca sua percepção apenas à trilha visual e sonora do audiovisual. Assim sendo, quando falávamos dos vídeos que exibem uma performance ao vivo de 4'33", o espectador do vídeo acaba por se tornar um meta-sujeito, por estar assistindo uma exibição realizada para outros espectadores – os sons que o espectador percebe serão os produzidos por aquele público no vídeo, e não os do ambiente em que ele assiste ao vídeo. O silêncio ainda é abstração, pois em cada vídeo ouve-se o que o microfone está captando: o motor do ônibus, o ar condicionado, os sons do público, e até mesmo o som da câmera de vídeo em funcionamento, numa metáfora digital dos sons do corpo humano (a câmera vaza o som dos circuitos eletrônicos e da fita rolando, o corpo humano vaza os sons do sistema nervoso e o coração batendo). Já nas versões eletrônicas de 4'33", essas sim contendo sonoramente o silêncio do audiovisual, o espectador está diante de um silêncio muito mais poderoso que o silêncio de um músico executando à sua frente a peça: é um silêncio *imaginado* e não *por abstração*.

Mas é curioso, em contrapartida, que esse silêncio digital acaba por afastar ainda mais esses vídeos da proposta de Cage: ele nunca procurou o silêncio absoluto com essa peça, pelo contrário: quis sempre mostrar o quanto o mundo está produzindo sons o tempo todo.

De qualquer forma, o que o audiovisual faz ao apropriar-se dessa irritação que a Música provoca nele, de ter de produzir silêncio na trilha visual e sonora, é, nos seus termos, produzir uma imagem-música: o silêncio de superfície. Não só a superfície da tela, mas a superfície dos discos rígidos contendo os códigos binários. Atendo-nos àquela versão em que a tela fica toda preta e a trilha sonora não produz som algum, o audiovisual fica completamente desprovido de informação: não há diferenças visuais, apenas repetição: *pixel* preto, *pixel* preto, *pixel* preto, *pixel* preto, *pixel* preto..., sendo que um *pixel* preto é formado por três pontos de luz, um azul, um vermelho e um verde, mas que são solicitados a ficarem apagados: então eles não são atualizados trinta vezes por segundo, como seria um *pixel* branco ou de qualquer outra cor, mas simplesmente não recebem informação alguma. Já na trilha sonora, como não há informação em bits, não há nada para ser convertido em pulso elétrico para vibrar as caixas acústicas. Na trilha sonora digital, há apenas zeros.

Por isso, o silêncio audiovisual é pura imaginação de superfície: enquanto o silêncio da nossa relação com o mundo é de pura diferença, a do audiovisual é pura repetição. Em compensação, ambos os casos provocam uma profunda irritação no espectador, mesmo que desprovidas de informação. É com essa concepção do silêncio imaginada pelo audiovisual que

encerro as análises: a última imagem-música reconhecida por esta dissertação.

Refletindo as descobertas provindas dos vídeos de 4'33" nos outros quatro tipos de vídeos musicais para *web*, é interessante observar a irônica função da repetição nestes vídeos: produzir diferenças. O exemplo que explicitarei no parágrafo anterior, do silêncio do audiovisual, é apenas um. Se voltarmos aos outros quatro tipos de vídeos analisados, podemos notar a importância da repetição em cada um.

Nos *Mashups* audiovisuais, em que a trilha sonora é ocupada por uma música pré-concebida, os processos de ressignificação têm necessariamente de ser propostos pelos compositores do vídeo a partir do que lançarão na trilha visual para se relacionar com a sonora. Quando uma imagem visual é utilizada para acompanhar um verso musical, a tendência é que se repita a mesma imagem se o verso se repetir na canção, e essa repetição é para potencializar a sincronia proposta: no vídeo *System Of A Vila – Chaves Suey*, poder ver por uma segunda vez Chiquinha gritando “*wake up!*” sendo respondida por Quico sussurrando “*wake up*” é o que potencializa e legitima tal sincronização. A cada repetição vai-se sedimentando a ideia de que são os personagens do Chaves que estão cantando tal canção, e não a banda System Of A Down. Mas essa ideia só pode existir como possibilidade, através da percepção do espectador, que sintetiza uma imagem-música ao somar, através de sua percepção, as imagens provindas da trilha visual e da sonora. Há uma performance que é apenas imaginada tecnicamente.

O mesmo efeito pode ser percebido nos vídeos compostos de *samplings* audiovisuais, mas nesse caso a ênfase está menos na sincronia, pois trilha visual e trilha sonora provêm de uma mesma fonte, e mais na diacronia, na capacidade de criar frases musicais a partir de elementos mínimos que são, em um só tempo, visuais e sonoros. A técnica mais recorrente é a de se criar frases melódicas ou rítmicas através da soma de diferentes samplers e os colocar em *loop*, criando territórios de significação por repetição típicos da música pop.

Não se encontra muita diferença no modo como os vídeos de *auto-tunning* operam: o método de montagem é praticamente o mesmo dos vídeos de *sampling*, com a adição de que a trilha de áudio passa pelo software Auto-Tune. O que muda no tratamento é justamente o uso desse software. Em alguns momentos em que se repete um *sampler* audiovisual, como, por exemplo, nos *loops* mencionados no parágrafo anterior, a melodia é modificada através desse software, o que produz uma diferença na repetição dos *samplers*.

Os *spoofs* de *shreds* são um caso a parte, pois não operam com essa lógica de fracionamento de imagens. No entanto, após as análises de 4'33", pode-se perceber que o espectador também se transforma numa espécie de meta-sujeito. O que ele observa não é uma



mera performance, mas sim a observação do usuário do *YouTube* às características visuais dessa performance, que ele mimetiza na trilha sonora.

Sintetizando o que foi apresentado nesse capítulo: evidenciou-se que novos modos de lidar com a relação entre Música e audiovisual estão se desenvolvendo em vídeos para *web*, que já não respondem às lógicas e à estrutura do videoclipe tipicamente televisivo: um novo ponto de observação precisa ser construído para que possa se compreender no que esses vídeos diferenciam-se de outros audiovisuais.

Foram reconhecidas algumas imagens-música que provocam a desterritorialização da Música através da linguagem audiovisual, o que em alguns momentos ocorre na destituição ou ressignificação de sentidos previamente arranjados pelas imagens de arquivo recicladas (que são, no caso, os vídeos não-musicais pré-publicados e que são utilizados como matéria prima para a composição dos vídeos musicais). Esse processo de atualização movimenta as estruturas da Música, forçando-a a criar novas informações (virtualização das imagens-música).

As quatro classes de vídeos musicais aqui apresentadas demonstraram como a linguagem, técnicas e estéticas do videoclipe foram importantes para se compreender os primeiros vídeos musicais para *web* que apresentavam algo de diferente em sua composição, como é o caso dos *mashups* e *sampling* audiovisuais, e que, no entanto, a tendência dos vídeos para *web* é a de se autonomizar, buscando uma linguagem que lhe é própria, caso dos *spoofs* de *shreds* e vídeos de *auto-tunning*.

Por fim, espera-se que tenha ficado clara a capacidade que a Música, ao sobrecodificar sua linguagem ao audiovisual, fez com que estes vídeos adquirissem: imaginar música, sintetizando imagens técnicas.

## 6 CONSIDERAÇÕES FINAIS

O principal desafio desta dissertação foi estudar a relação entre audiovisual e Música através de um olhar que privilegiasse a musicalidade dessa relação, pensando a Música como capaz de sobrecodificar a linguagem audiovisual, invertendo o ponto de vista que normalmente se lança a outros audiovisuais, principalmente ao videoclipe. Os vídeos para *web* aqui estudados estabelecem outro paradigma, no qual a virtualidade da Música é que vai estabelecer os processos de significação. Talvez só tenha sido possível perceber esse movimento através da teoria de Luhmann, através da qual se trocou o ponto de vista sujeito/objeto para o de sistema/meio. Assim, pôde-se perceber que nos vídeos analisados a Música deixa de ter uma característica de subordinação a uma narrativa, passando a agir, a interferir, a irritar o audiovisual. Talvez pensando através deste ponto de vista, possa se visitar alguns videoclipes e filmes para tentar encontrar essa ordem de ação da Música.

Através das análises, pôde-se perceber que de fato a Música imprime algo de si nos vídeos musicais para *web*, e também se diferencia de si ao se territorializar neste meio. Mas esse diferenciar-se de si não resulta em uma nova música no sentido que normalmente a compreendemos, como uma canção com estruturas diferentes ou um novo estilo; não está tampouco atrelado a sonoridades. Não são canções novas, mas novos procedimentos composicionais que se atualizam nestes vídeos. Além disso, elas são capazes de apresentar a possibilidade de manifestações musicais que não sejam estritamente sonoras, mas visuais, corporais, multissensoriais. O que a imagem-música faz é apresentar possíveis, não da mesma forma como as notas desenhadas na partitura o fazem, mas com a mesma potência de se realizarem: elas não representam sons no mundo, mas conceitos de sons, conceitos de Música – por isso que não precisam ser estritamente sonoros.

Ao pensar na virtualidade da Música, ela é observada aqui na maior parte do tempo como música pop. Desse modo, para a música pop, a performance, a coreografia, a moda, tantos outros elementos também *são* Música. Não dá para pensá-las como da linguagem musical, mas como musicalidades.

Isso pôde ser percebido porque adotei uma perspectiva bergsoniana: se o universo todo é composto por imagens, perde o sentido dividir o audiovisual em *imagem* e *som*, bem como pensar na imagem como algo estritamente visual e na música como algo estritamente sonoro. A trilha visual transmite – e também provoca na percepção – imagens visuais, bem como a trilha sonora, imagens sonoras, e qualquer imagem pode adquirir características musicais.

Esse foi um passo importante para poder pensar a Música além da trilha sonora, reconheço ao se manifestar em imagens-música, tanto as possíveis, que expandem seu virtual, quanto as atuais, que resolvem uma problemática desse mesmo virtual. Penso que a atualização do conceito de imagem-música, feita por essa dissertação, possa permitir outros modos de observação e análise de audiovisuais musicais (ou da Música no audiovisual), encontrando imagens não apenas no plano do realizado, mas também no do possível.

Ainda sobre os resultados da análise, inicialmente, a proposta da dissertação era estudar os vídeos de modo a restituir de algum modo o movimento deles textualmente. No entanto, isso só se fez necessário nos vídeos compostos de *samplers* audiovisuais (feito através de sequências de *frames* e gráficos), para poder perceber como eles produziam sentido nas articulações entre *samplers* diacronicamente. Já os *mashups*, *spoofs* de *shreds* e *auto-tunnings*, ainda que impliquem a compreensão dos *samplers* para serem entendidos, apresentaram outras manifestações potenciais da Música, que se dão a ver de outras formas, como no choque entre imagem-percepção e imagem-lembrança provocada pelo vídeo *John Williams Shreds*; ou nos *mashups*, que imaginam uma performance impossível de ser realizada fisicamente através da articulação de trilhas visuais e sonoras que não possuíam relação alguma entre si; ou ainda na possibilidade de fazer com que âncoras de telejornal e entrevistados se tornem capazes de cantar, através do Auto-Tune; e até mesmo de se desenvolver uma imagem do silêncio audiovisual.

A problematização da relação entre imagem e Música em audiovisuais, no primeiro capítulo, ajudou a perceber que os estudos sobre videoclipe não dariam conta de responder às questões que foram elaboradas a partir da observação empírica do corpus. Ainda que os videoclipes estejam encontrando espaço na *web*, eles continuam sendo pensados para serem transmitidos nas mais diferentes mídias, enquanto que para boa parte dos usuários que publicam vídeos musicais no *YouTube* esta não é uma preocupação. Embora vídeos como os *VideoSongs* possuam um formato mais passível de ser transmitido na televisão, por não infringir direitos autorais, apresentar uma canção popular nos moldes da indústria fonográfica e inserir na trilha visual a imagem dos músicos, ainda assim o formato como são compostos não encontra paralelos nos videoclipes que vemos na televisão: eles derivam de outros métodos de montagem que fizeram sucesso dentro do próprio *YouTube*: a remixagem, o uso de *loops*, sequências de *samplers* audiovisuais em repetição incessante, a reciclagem de material previamente publicado, e assim por diante.

O que a maior parte dos vídeos aqui analisados faz é repetir *samplers* audiovisuais para criar *loops*, que são típicos da música pop, e que acabam por criar territórios: se a

montagem, a simultaneidade de várias molduras na tela e os cortes velocíssimos de *sampling* audiovisual podem deixar o espectador “à deriva”, a repetição em *loop* de boa parte desses mesmos *samplers* põe o espectador numa zona de conforto, num ciclo de tempos espacializados que lhe dá a possibilidade de perceber o tempo de duração de cada *sampler*. A repetição (às vezes excessiva) das notas, das frases e das estruturas musicais são fundamentais para a música pop, que não opera em “profundidade”, mas em superfície. Uma canção popular moderna não é auto-suficiente em processos de significação, ao menos não do mesmo modo como uma fuga de Bach, o qual poderíamos passar anos a fio analisando as relações melódicas. A música pop, por sua vez, se expande para fora de si, efetuando semiose em outros elementos que não necessariamente são sonoros: coreografia, performance, moda, comportamento, tribos musicais, cinema, televisão, internet... a música pop não só significa, como também virtualiza-se nessas esferas, criando imagens-música. Ela opera expandindo sua superfície. Há menos a ser interpretado nela mesma, por isso que ela se repete incessantemente. Importa agora explorar a vasta superfície que a música popular territorializa, rizomaticamente, para além de seus domínios sonoros. As práticas de montagem visual dos vídeos musicais para *web* são expressão dessa virtualização para além da linguagem musical e da sua manifestação sonora: os *loops*, as repetições, o *sampling*, a conversão em melodia de eventos sonoros, etc.

O pensamento em superfície de Flusser não dá atenção necessariamente à bidimensionalidade das telas, mas refere-se também à superficialidade das relações de diferentes suportes e linguagens, de um achatamento que está dirigindo as linguagens à nulodimensionalidade, à produção de mundos através de cálculos computacionais. É nesse cenário que se torna, enfim, possível, desenvolver música imaginativa, *música de superfície*.

Espero que esta pesquisa tenha contribuído no entendimento desses fenômenos que estão ocorrendo nos vídeos para *web*, um campo que ainda estamos aprendendo a explorar. Também espero que tenha contribuído, principalmente, para um novo modo de observar a relação entre Música e audiovisual. Observando-a como uma virtualidade que irrita o audiovisual, as distinções e comparações entre o visual e o sonoro deixam de ser o centro das atenções, e novas problemáticas surgem para a pesquisa em audiovisuais musicais.

## REFERÊNCIAS BIBLIOGRÁFICAS

AUMONT, Jaques; MARIE, Michel. **Dicionário teórico e crítico de cinema**. Campinas: Papyrus, 2003.

BERGSON, Henri. **Matéria e Memória**. São Paulo: Martins Fontes, 1990.

\_\_\_\_\_. **A evolução criadora**. São Paulo: Martins Fontes, 2005.

\_\_\_\_\_. **Duração e simultaneidade**. São Paulo: Martins Fontes, 2006a.

\_\_\_\_\_. **O pensamento e o movente**. São Paulo: Martins Fontes, 2006b.

\_\_\_\_\_. **Memória e vida**. São Paulo: Martins Fontes, 2006c.

BJÖRNBERG, Alf. Structural relationships of music and images in music video. In: MIDDLETON, Richard. **Reading pop: approaches to textual analysis in popular music**. Nova Iorque: Oxford, 2000, pp. 346-378.

BORNHEIM, Gerd. **Metafísica e finitude**. São Paulo: Perspectiva, 2001.

CAGE, John. Transcrição de entrevista concedida no filme “**Ecoute**” por Miroslav Sebestik, 1991. Disponível em: <<http://hearingvoices.com/news/2009/09/cage-silence/>>. Acesso em: 22 jul. 2010.

CARVALHO, Any Raquel. **Contraponto modal: manual prático**. 2. ed. Porto Alegre: Evangraf, 2006.

CHAGAS, Paulo C. A Música de Câmara Telemática: a Metáfora de Flusser e o Universo da Música Eletroacústica. In: **GHREBH: Comunicação, Imagem e Técnica - Vilém Flusser**. v. 1, n. 11. Cisc: São Paulo, 2008. Disponível em: <<http://www.cisc.org.br/revista/ghrebh/index.php?journal=ghrebh&page=article&op=view&path%5B%5D=7&path%5B%5D=5>>. Acesso em: 1 mar. 2011.

CHION, Michel. **A audiovisualização: som e imagem no cinema**. Lisboa: Texto & Grafia, 2008.

COELHO, Lillian. As relações entre canção, imagem e narrativa nos videoclipes. In:

INTERCOM - XXVI Congresso Brasileiro de Ciências da Comunicação, Belo Horizonte. **Anais**, 2003. Disponível em: <<http://www.ericaribeiro.com/Arquivos/CancaoImagemNarrativaVideoclipes.pdf>>. Acesso em: 11 out. 2010.

COELHO NETTO, José Teixeira. **Moderno pós moderno**. São Paulo: Iluminuras, 1995.

CONTER, Marcelo B.; KILPP, Suzana. Novos construtos de tempos audiovisuais simultâneos no videoclipe. In: **Comunicação, Mídia e Consumo** (São Paulo. Impresso), v. 6, p. 181-202, 2010.

COSTA, J. P. Fluxos videomusicais: uma proposta epistemológica para o estudo dos vídeos musicais na *web* 2.0. In: **PRISMA.COM**, n. 9, 2010.

DELEUZE, Gilles. **A imagem-movimento**. São Paulo: Brasiliense, 1985.

\_\_\_\_\_. **Bergsonismo**. São Paulo: Editora 34, 2004.

\_\_\_\_\_. **A Ilha Deserta**. São Paulo: Iluminuras, 2005.

\_\_\_\_\_. **Lógica do Sentido**. São Paulo: Perspectiva, 2007.

\_\_\_\_\_. **A imagem-tempo**. São Paulo: Brasiliense, 2009.

\_\_\_\_\_; GUATTARI, Félix. **Mil Platôs - capitalismo e esquizofrenia, vol. 4**. São Paulo: Ed. 34, 1997.

ECO, Umberto. **Obra Aberta**. São Paulo: Perspectiva, 2007.

EISENSTEIN, Serguei. **O sentido do filme**. Rio de Janeiro: Jorge Zahar, 2002.

FELINTO, Erick. Videotrash: o YouTube e a cultura do “spoof” na internet. **Revista Galáxia**, São Paulo, n. 16, p. 33-42, dez. 2008.

FERNÁNDEZ SANCHEZ, Manuel Carlos. **Influencia del montaje en la lenguaje audiovisual**. Madrid: Libertarias, 1997.

FLUSSER, Vilém. **O mundo codificado**: por uma filosofia do design e da comunicação. São Paulo: Cosac Naify, 2007.

\_\_\_\_\_. **O universo das imagens técnicas**: elogio da superficialidade. São Paulo: Annablume, 2008.

\_\_\_\_\_. **Filosofia da caixa preta**: ensaios para uma futura filosofia da fotografia. São Paulo: Annablume, 2011a.

\_\_\_\_\_. **Pós-história**. Vinte instantâneos e um modo de usar. São Paulo: Annablume, 2011b.

FORZANARI, Sandro Kobol. A imagem-cristal: a leitura deleuziana de Bergson nos livros sobre o cinema. IN: **Artefilosofia**. n. 9, Ouro Preto: UFOP, 2010. p. 93-100.

FRITH, Simon. '**Making sense on music video**: pop into the nineties'. In: **Music for pleasure**. Cambridge, 1988. pp. 205-225.

GOODWIN, Andrew. **Dancing in the distraction factory** – music television and popular culture. Minneapolis: University of Minnesota Press, 1992.

HOLZBACH, Ariane D. ; NARCOLINI, Marildo José. Videoclipe em tempo de reconfigurações. IN: **Revista FAMECOS** (Online), v. 1, p. 50-56, 2009.

ILHARCO, Fernando. **Filosofia da Informação**: Uma Introdução à informação como fundação da acção e da comunicação e da decisão. Lisboa: Universidade Católica, 2003, p. 133-185.

LÉVY, Pierre. **O que é o virtual?**. São Paulo: Ed. 34, 2005.

LUHMANN, Nicklas. **Introdução à teoria dos sistemas**. Petrópolis: Vozes, 2009.

MACHADO, Arlindo. Anamorfoses cronotópicas ou a quarta dimensão da imagem. In PARENTE, André (Org.). **Imagem-Máquina**: A era das tecnologias do virtual. Rio de Janeiro: Editora 34, 1993.

\_\_\_\_\_. **A televisão levada a sério**. São Paulo: SENAC, 2009.

MANZANO, Luiz Adelmo Fernandes. **Som-imagem no cinema**: a experiência alemã de Fritz Lang. São Paulo: Perspectiva, 2003.

MARCONDES FILHO, Ciro. **O escavador de silêncios**: formas de construir e de desconstruir sentidos na comunicação. São Paulo: Paulus, 2004.

NASCIMENTO, Guilherme. **Música menor**: a *avant-garde* e as manifestações menores na música contemporânea. Annablume: São Paulo, 2005.

NÖTH, Winfried; SANTAELLA, Lúcia. **Imagem**: Cognição, Semiótica, Mídia. São Paulo: Iluminuras, 2009.

PEIXOTO, Nelson Brissac. Passagens da imagem: pintura, fotografia, cinema, arquitetura. In: PARENTE, André (Org.). **Imagem-Máquina**: A era das tecnologias do virtual. Rio de Janeiro: Editora 34, 1993.

PRENDERGAST, Roy M. **Film music**: a neglected art: a study of music in films. 2ª ed. New York: W. W. Norton, 1992.

SÁ, Simone Pereira de. A trilha sonora de uma história silenciosa: som, música, audibilidades e tecnologias na perspectiva dos Estudo de Som. In: **Rumos da cultura da música**: negócios, estéticas, linguagens e audibilidades. Organizado por Simone Pereira de Sá. Porto Alegre: Sulina, 2010. pp. 91-109.

SADIE, Stanley. Dicionário grove de música: edição concisa. Rio de Janeiro: Jorge Zahar, 1994.

SANTAELLA, Lúcia. **Matrizes da linguagem e pensamento**: sonora visual verbal: aplicações na hipermídia. São Paulo: Iluminuras, 2005.

SAUSSURE, Ferdinand de. **Curso de lingüística geral**. São Paulo: Cultrix, 1993.

SCHAFER, R. Murray. **O ouvido pensante**. São Paulo: UNESP, 1991.

\_\_\_\_\_. **A afinação do mundo**: uma exploração pioneira pela história passada e pelo atual estado do mais negligenciado aspecto do nosso ambiente: a paisagem sonora. São Paulo: UNESP, 2001.



SILVA, A. R. A imagem-música. In: **Comunicação: Veredas** (UNIMAR), v. 5, p. 39-49, 2006. Disponível em: <<http://www.unimar.br/publicacoes/comunicacao05.pdf>>. Acesso: 26 out. 2010.

SILVA, A. R.; ROSSINI, M. S (orgs.). **Do audiovisual às audiovisuais**: convergência e dispersão nas mídias. Porto Alegre: Asterisco, 2009.

SOARES, Thiago. **Videoclipe**: o elogio da desarmonia. Recife: LivroRápido, 2004.

\_\_\_\_\_. Videoclipe, You Tube e Televisão Expandida: Notas sobre Itinerários do Audiovisual na Cultura Midiática. In: XXXII Congresso Brasileiro de Ciências da Comunicação, Curitiba. **Anais...**, 2009. Disponível em: <<http://www.intercom.org.br/papers/nacionais/2009/resumos/R4-0863-1.pdf>>. Acesso: 6 jun. 2011.

STERNE, Jonathan. O MP3 como um artefato cultural. In: **Rumos da cultura da música**: negócios, estéticas, linguagens e audibilidades. Organizado por Simone Pereira de Sá. Porto Alegre: Sulina, 2010. pp. 63-89.

VERNALLIS, Carol. **Experiencing Music Video**: Aesthetics and Cultural Context. New York: Columbia, 2004.

\_\_\_\_\_. Music video, songs, sound: experience, technique and emotion in Eternal Sunshine of the Spotless Mind. IN: **Screen**, vol. 49, nº 3, Oxford: Oxford University Press, 2008.

WISNIK, José Miguel. **O som e o sentido**: uma outra história das músicas. São Paulo: Companhia das letras, 1999.

## GLOSSÁRIO

A Capella - Música vocal sem acompanhamento instrumental.

Contraponto - “[...] é a arte de combinar duas ou mais linhas melódicas independentes e simultâneas, de acordo com uma série de regras” (CARVALHO, 2006, p. 11).

Feedback - Em música, ocorre quando um microfone captura os sons que ele mesmo está enviando a um alto falante, provocando um ruído interminável.

Frame - Embora a tela de televisores e monitores seja constituída por uma malha de pixels, de pontos luminosos isolados, é possível, através de softwares, compor uma imagem congelada de uma varredura completa da tela, que é compreendida por *frame*. O tempo de varredura do vídeo para compor um *frame* corresponde a 1/30 de segundo.

Harmonia - Sucessão de acordes, em geral com uma linha melódica mais importante (CARVALHO, 2006, p. 50).

Leitmotif - Em música, o termo designa um tema melódico utilizado para caracterizar um personagem na ópera. O cinema mais tarde se apropriaria dessa ideia, mas nos seus termos, como é o caso do tema do filme *Tubarão* (Steven Spielberg, 1975).

Loop - Técnica de gravação fonográfica que consiste, quando o material é fita, em elaborar anéis de fita, recortados de uma fita normal e coladas as suas pontas. Os softwares simulam esse efeito colocando em repetição qualquer amostra sonora que se dejesar.

MC - Abreviação de Mestre de Cerimônias, atribui-se ao cantor de rap, hip hop, funk carioca e outros estilos.

On the beat - Técnica de sincronização de imagens visuais e sonoras. Em videoclipes, trata-se fundamentalmente de trocar de plano visual no mesmo momento em que a música muda de compasso.

Pitch shifting - Alteração da altura melódica de um evento sonoro, tornando-o mais grave ou agudo.

Playback - Procedimento técnico em que uma música pré-gravada é aplicada à trilha de áudio, enquanto na trilha da imagem ela é dublada por músicos.

Plugin - É um espécie de *software*, mas um programa menor, com uma finalidade muito específica, e que é instalado como extensão de um *software* maior.

Riff - Conjunto de notas ou acordes musicais que servem de acompanhamento de canções. Um exemplo notável é o riff de guitarra que inicia a canção *Satisfaction* da banda The Rolling Stones.

Ritardando - Em música, é o retardamento gradual do andamento da música, fazendo com que os compassos sejam progressivamente mais longos.

Scratching - Interferir na leitura de um disco de vinil manualmente, como os mestres de cerimônia de rap o fazem. O *scratching* foi adaptado pelos softwares de edição de áudio e vídeo, nos quais é possível simular o efeito.

Staccato - Ataque seco das notas musicais, fazendo com que haja um intervalo de silêncio entre uma nota e outra.

Stone rock - Sub-gênero do *Grunge*, cujas bandas de destaque são Nirvana, Pearl Jam e Mudhoney, o *Stone rock* tem como algumas de suas características que se diferenciam do *grunge* timbres ainda mais sujos e andamento mais rápido.