

A imagem-música

A configuração do conceito de imagem-música remete, inicialmente, a dois estudos apresentados por Gilles Deleuze nos anos 80 relativos ao cinema: *A imagem-movimento* (1984) e *A imagem-tempo* (1990), em que o filósofo francês, ao estabelecer uma cartografia das imagens cinematográficas, identificava, respectivamente, (1) uma tipologia de imagens - imagem-afecção, imagem-ação e imagem-percepção -, inspirado pela obra de Henri Bergson e de Charles Sanders Peirce -, e (2) o conceito de zeroidade, a partir do qual problematizava a questão do tempo e reivindicava uma categoria, menos descritiva do que aquelas apresentadas pela semiologia e mesmo por Peirce, capaz de expressar a natureza do devir das imagens.

Tal percurso, associado à evidência de que um modo idêntico de produção de imagens e de músicas se apresenta a partir dos meios digitais, cuja característica fundamental consiste em partir de modelos altamente abstratos, matemáticos, independentes de seus referentes ou de uma subjetividade individual, conduziu a elaboração do conceito de imagem-música a ser desenvolvido neste artigo.

Entre o final do século XIX e o início do século XX, o filósofo Charles Sanders Peirce propunha à discussão a hipótese de que a tendência de todas as ciências era aumentar o seu grau de abstração até se saturarem na matemática. Santaella e Nöth (1998), no contexto de uma semiótica aplicada, perceberam que tal movimento também se verifica quando o objeto de estudo são as linguagens:

parece haver uma tendência atual, em todas as linguagens, de caminharem para um modo de estruturação, para uma morfogênese semelhante à da música, que sempre foi, aliás, no seu movimento de puras relações, a mais matemática dentre todas as linguagens. (NÖTH, SANTAELLA, 1998, p. 90)

A expressão de tal tendência pode ser encontrada naquilo a que se denomina, aqui, imagem-música. Imagem-música é um texto sincrético, isto 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.

Tal constatação implica reconhecer que o problema das audiovisualidades teve um deslocamento: do estudo das múltiplas linguagens que as formam (som e imagem), com suas gramáticas específicas, para a descrição de uma nova gramaticalidade capaz de demonstrar os modos de produção de quaisquer textos audiovisuais, no caso, de quaisquer imagens-músicas.

Arlindo Machado reforça tal idéia ao dizer que não se trata mais de demonstrar as semelhanças (e, pode-se acrescentar aqui, mesmo as diferenças) entre imagens eletrônicas e sons eletrônicos, buscando analogias entre sons e imagens, mas de reconhecer a existência de uma outra máquina semiótica produtora de signos que engendra textos sincréticos.

Esta máquina, segundo Nöth e Santaella depende “de programas, de valores numéricos, e de procedimentos específicos, algoritmos de simulação do som ou da imagem, para serem transmitidos nos terminais específicos de efeito sensível para o olho ou para o ouvido.” (1998, p. 91). E seu modo de operação, sua gramática, continuam, assemelham-se às gramáticas musicais:

(...) 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. (NÖTH, SANTAELLA, 1998, p. 91)

Ora, se não há diferenças no modo de formar sons (musicais) e imagens, e se este modo é passível de uma descrição formal e não-contraditória, está-se diante de um sistema semiótico cujos elementos articulam-se a partir de regras estruturais cuja natureza parece estar próxima à matemática e à música, conforme sugerem Santaella, Nöth e Machado.

Assim,

(...) a chave semiótica da computação gráfica não está só na imagem, mas nas ligações indissolúveis da imagem computacional com a forma de engendramento que é constitutiva da sintaxe sonora (...) A questão do

tempo como passagem, sucessividade, evanescência, não pertence à lógica da visualidade, mas sim à lógica da narrativa, que é eminentemente verbal, e muito mais especialmente à lógica da música, que se constitui no território onde o tempo reina soberano.” (NÖTH, SANTAELLA, 1998, p.89).

Esse problema da temporalidade, bem como sua conceituação, acompanha o desenvolvimento das mídias, contudo não fica restrito a elas e a suas gramáticas. Tal ressalva, a ser discutida a seguir, não impede que se identifiquem nos processos históricos das mídias, modalidades de expressão sógnicas do tempo.

O signo, para Peirce, sempre representa alguma coisa para alguém, desde uma determinada condição, desde um meio ou representâmen, desde uma mídia, pode-se acrescentar. A natureza deste meio, portanto, ao mesmo tempo em que estabelece os limites da representação também engendra mundos derivados de suas potencialidades específicas. Dessa forma, o tempo engendrado por um veículo capaz de produzir imagens fixas, como a fotografia, por exemplo, só pode ser um tempo representado. A existência do tempo, aqui, é representação. No cinema, tem-se a ilusão do tempo, de acordo com Santaella e Nöth (1998, p. 93), ou a criação de imagens-movimento e imagens-tempo, conforme nos ensina Gilles Deleuze (1984; 1990) em seus escritos sobre cinema. Esta criação é maquinação, arranjo, agenciamento de desejo e de enunciação. No entanto, em sua relação com o objeto, o cinema, por sua natureza sógnica, apresenta o tempo como ilusão de movimento.

É com o vídeo que entramos no “universo das imagens em movimento-tempo real” (SANTAELLA; NÖTH, 1998, p. 93). Conforme Arlindo Machado (1988, p. 43), “o vídeo, por consequência de sua própria constituição, é o primeiro mídia a trabalhar concretamente com o movimento, isto é, com a relação espaço-tempo”. A imagem no vídeo

que as lentes refratam é projetada numa superfície foto-sensível (target), cuja capacidade para conduzir eletricidade varia de acordo com a quantidade de luz que incide sobre cada um de seus pontos. Dessa forma, ao ser varrida pelo feixe de elétrons, a luminosidade da imagem em qualquer ponto é traduzida em amplitude de um sinal elétrico, de modo que cada ponto ou retícula do espaço bidimensional é convertido em nível de voltagem de um impulso elétrico na seqüência temporal. À medida que a intensidade luminosa da imagem varia de um ponto ao outro da linha de varredura, a amplitude do sinal se modifica de forma sincronizada. Assim, uma imagem projetada no suporte fotocondutor é traduzida em mudanças na voltagem de um sinal elétrico durante o tempo necessário para fazer o seu esquadramento completo (MACHADO APUD SANTAELLA NÖTH, p. 77)

No entanto, apenas com o computador é que o tempo passou a ser introjetado na imagem, “imagem que, finalmente, adquire o poder de se comportar exatamente como o som na sua natureza de puro tempo.” (NÖTH, SANTAELLA, 1998, p. 93). Assim, torna-se plausível que a busca da gramaticalidade audiovisual passe pela descrição desta nova máquina semiótica produtora de signos, que é o computador, sobretudo em seus aspectos genuinamente matemáticos, e pela recuperação das reflexões já feitas sobre a música que, também em Nietzsche, desempenhava um papel fundamental para a compreensão dos modos como se produzem sentidos.

Não se pode, entretanto, negligenciar o segundo aspecto referido por Santaella e Nöth quando dizem que, mesmo que não haja diferenciação nos modos (gramaticais) de formação da imagem-música, as diferenças surgem nos seus modos de aparição, na maneira como se apresentam para os sentidos.

A tensão produzida pelos signos na vida social sobre as gramáticas revela a ação positiva desempenhada pelas micropolíticas (pós-midiáticas)¹ que, por sua vez, só são reconhecidas *à posteriori*, quando novas mídias e tecnologias se impõem (DERRIDA, *Ecografías de la televisión*, 1998). Esta dimensão de tensão, denominada como agramaticalidade, é uma noção filosófica que “põe em jogo o funcionamento positivo, sobretudo intensivo, da linguagem; não uma apreciação sobre modos ou momentos em que a linguagem perde algo de si, deixa de ser possível [...] o agramatical em Deleuze é sempre um real, embora seu modo de ser seja o da virtualidade, em que se agitam potências ditas de diferenciação.” (ALMEIDA, 2003, p. 29-30). Deleuze complementar, em *Crítica e Clínica*:

Será possível fazer gaguejar a língua sem confundi-la com a fala? Tudo depende, na verdade, da maneira pela qual se considera a língua: se a tomamos como um sistema homogêneo em equilíbrio [a gramática computacional-matemática da imagem-música], ou próximo do equilíbrio, definido por termos e relações constantes, é evidente que os

¹ O conceito de pós-mídia é desenvolvido em *Cartographies schizoanalytiques* (GUATTARI, 1989) e em *Elementos para uma comunicação pós-midiática* (SILVA, 2003).

desequilíbrios e as variações só afetarão as palavras (variações não-pertinentes do tipo entonação[...]). Mas se o sistema se apresenta em desequilíbrio perpétuo, em bifurcação, com termos que, por sua vez percorrem, cada qual, uma zona de variação contínua, então a própria língua põe-se a vibrar, a gaguejar, sem contudo confundir-se com a fala” (DELEUZE, 1997, p. 123)

Aqui aparecem relevantes os estudos relativos a gêneros e formatos televisivos não para descrever suas gramáticas próprias apenas, mas, sobretudo, para observar como gêneros e formatos podem gaguejar², entrar em variação contínua em relação a outros gêneros e formatos e, simultaneamente, como podem ser afetados pelas relações intersemióticas, no caso realizadas entre imagens e músicas, que imaginam poder expressar. Ou seja, há três níveis em que as conexões operam: o primeiro, realizado no interior de um dado programa, quando o gênero formata suas linguagens e estabelece previamente os limites de suas relações intersemióticas; o segundo, necessariamente híbrido, dual, quando uma música se faz representar em uma imagem ou o contrário; e o terceiro, caracterizado pelo princípio da simulação, quando os modelos realizam a diversidade de imagens e sons apresentados aos sentidos.

François Jost, em *Seis lições sobre televisão* (2004), ao discutir o caráter pactual do gênero, argumenta que “o problema não é apenas saber o que é um telejornal ou um documentário. É preciso saber, também, como a própria televisão denomina tal e tal objeto, como o define.” (JOST, 2004, p. 17) Essa observação de Jost aponta, simultaneamente, para o primeiro nível de conexão referida, dependente da natureza auto-referencial do meio e, também, para sua variabilidade, uma vez que, ao entrar em conexão com outros meios, pode fazer variar suas leis gerais e o modo como se auto-define, sempre a partir daquilo que Jost define como promessa.

Para o autor, a promessa, ao contrário do contrato ou do pacto, não define o gênero por um saber codificado, tal como o proclamava Umberto Eco, mas estabelece uma interface entre emissor (televisão) e telespectador, de forma a garantir aos emissores a possibilidade de ‘etiquetar’ seus programas – Festival de Música Popular Brasileira, O Fino da Bossa, Som Livre Exportação, por exemplo – para configurar a promessa que fazem de veicular programação musical e, aos telespectadores, a possibilidade de verificar se a promessa foi efetivada, razão pela qual “quando se estuda televisão, não se deve ficar restrito apenas à consideração do próprio programa, mas tem-se de estudar o que se fala a seu respeito, como se fala dele e o que se diz”. (JOST, 2004, p. 18). Abrem-se, assim, os estudos televisivos à problemática das conexões referidas por Deleuze, recuperando a dimensão de um nível propriamente micropolítico prévio às gramáticas e textualizações, uma vez que sempre haverá a possibilidade de se verificar se a promessa foi efetivamente cumprida.

Ainda sobre o primeiro nível de conexão, observou-se que o gênero conforma suas matérias para que se realizem as promessas feitas. É importante preservar esta dimensão em sua especificidade porque ela refere, sucessivamente e em grau de determinação decrescente, a natureza do meio, a identidade da emissora, o gênero de programação e o programa com suas estéticas, gramáticas e textualizações específicas. Cada um desses níveis possui uma gramaticalidade específica que se atualiza no nível referido e se re-atualiza nos subseqüentes, quando não ocorrem ruídos.

O movimento das sobre-determinações parece fundamental para que se observem como se vão reproduzindo, nos mundos televisivos³, as promessas do meio, mas também para que se percebam manifestações dispersivas em relação aos códigos que o caracterizam. Tais dispersões são expressas por manifestações híbridas que preparam as percepções para mudanças do meio em devir.

No Brasil, a primeira cantora propriamente produzida pela televisão - Elis Regina⁴ - rompe com a tradição vigente de música televisionada e passa a investir na articulação dessas duas linguagens: a imagem e a música. Segundo Augusto de Campos,

Elis [...] teve, realmente, um grande mérito no sentido da popularização da Bossa Nova [...]. Suas interpretações elétricas e eletrizantes, a alegria contagiosa que transmitia, não tanto com a sua voz (que nada tem de excepcional), mas com um compósito de voz e corpo, canto e coreografia articulados numa alegria juvenil e irresistível, explodiram como uma verdadeira bomba no samba, com um alto poder de comunicação. [...] Elis extroverteu a BN, desencravou-a, tirando-a do âmbito restrito de câmara e colocou-a no palco-auditório de TV.” (CAMPOS, 1993, p. 54)

² Gilles Deleuze discute o problema da gagueira em *Crítica e clínica* (1997, p. 122-9).

³ A expressão Mundos Televisivos foi proposta por Suzana Kilpp (2005), inspirada pelo artigo Mundos Artísticos e Tipos Sociais, de Howard S. Becker (VELHO, 1977, p. 9 – 26).

⁴ Elis Regina vence em 1965 o I Festival de Música Brasileira, na TV Excelsior, com a música Arrastão. Segundo Caetano Veloso, mesmo que Elis não fosse uma boa cantora, teria importância fundamental na cultura brasileira por ter sido a primeira cantora produzida pela TV

Assim, tem-se como perceber, no nível do programa e suas estéticas, simultaneamente uma textualização (voz e corpo) que realiza as determinações do meio, rompendo com a tradição que, na tevê, apresentava imagem e som paralelamente, propondo um híbrido de imagem e som. Ainda não se trata de imagem-música, mas se configura como uma etapa intermediária que, por um lado, realiza musicalmente o projeto do mundo televisivo e, por outro, antecipa a autonomização dos modelos computacionais que produzem imagem-música.

O desenvolvimento de tal projeto televisivo-musical teve seus pontos de crise, que apontam para o segundo nível de conexão: quando a música se faz imagem e o contrário. Augusto de Campos, preocupado com o projeto estético da Bossa Nova, e preso a parâmetros dela, não percebe a natureza da crise e a reduz a uma espécie de 'exageração', quando se refere à figura televisiva de Elis:

"[...] talvez pelo afã de ampliar o público, o programa [O fino da Bossa] foi-se tornando cada vez mais eclético, foi deixando de ser o porta-voz da BN para se converter numa antologia mais ou menos indiferente dos hits da música popular brasileira [...]. Por seu turno, a própria Elis foi sendo levada a uma exageração do estilo interpretativo que criara. Seus gestos foram-se tornando cada vez mais hieráticos. Os rictos faciais foram introduzidos com frequência sempre mais acentuada. A gesticulação, de expressiva passou a ser francamente expressionista. [...] É uma interpretação rígida, enfática, de efeitos melodramáticos (inclusive jogos fáceis de iluminação cênica). Esse estilo de interpretação 'teatral' quase nada mais tem a ver com o estilo de canto típico da BN." (CAMPOS, 1993, p. 55)

O canto proposto por Elis pouco tem a ver, efetivamente, nessa fase, com a Bossa Nova, mas tem muito a ver com o Mundo Televisivo que se insinuava fortemente. Talvez o próprio conceito de canto e seus parâmetros devessem aqui ser questionados. O que se forjava era uma espécie de canto-imagem, um devir de imagem-música, ainda não digitalizado.

O duelo entre música e voz, por um lado, e imagem e corpo televisivo, por outro, atualiza esse segundo nível de conexões. Quando a música se faz imagem? Quando a imagem se faz música? Que processo de hibridação foi esse que anteviu, por outros meios, a imagem-música que se realiza como potencialidade tecnológica depois da utilização do computador para tais fins e que configurou Elis Regina como voz-corpo-televisivo?

Antes de prosseguir essa discussão, abordando a imagem-música como terceiro nível de conexão sustentada pelo princípio da autonomia, fazem-se necessárias algumas considerações sobre o propriamente televisivo, como forma-televisiva, e sobre a cena a ser televisionada, cuja organização semiótica se estatui, também, a partir de outros princípios de formalização não-televisivos.

Ao abordar a televisão devem-se considerar metodologicamente sobretudo as formas televisivas que produzem os Mundos Televisivos e suas cenas, de modo que se criem, no interior desses Mundos, sujeitos televisivos, objetos televisivos e relações televisivas. O primado da forma sobre suas matérias circunscreve, na perspectiva aqui adotada, o próprio de uma dada semiótica, ou seja, o próprio da televisão.

Hjelmslev (1961) reconhece a existência de um mundo extra-semiótico e defende que abordagens semióticas e não-semióticas devam ser feitas independentemente uma da outra, de tal modo que as formas produzidas pelas disciplinas não-semióticas sirvam, em outro contexto, como substâncias de novas formas semióticas. Isso implica reconhecer que a formação semiótica "se baseia não no sentido, mas no próprio princípio da forma e nas possibilidades que decorrem de sua realização" (Hjelmslev, 1961, p. 80).

Assim, o propriamente televisivo, do ponto de vista da semiótica, é identificável na descrição das formas televisivas e não na consideração das cenas que são televisionadas independentemente. As cenas televisionadas são substâncias da forma televisiva. Portanto, o que tais cenas fazem é evidenciar novas possibilidades decorrentes da realização da forma. É sob este aspecto que Elis Regina, como cena televisiva, tensionou as gramáticas formais no sentido de criar novas realizações, algumas delas expressas posteriormente pelos vídeos e outras, mais autônomas, expressas como imagem-música⁵.

O *Primeiro Festival de Música Brasileira*, exibido pela TV Excelsior em 1965 e que popularizou nacionalmente Elis Regina, assim como *O Fino da Bossa*, programa comandado por Jair Rodrigues e pela cantora de 1965 a 1967 na TV Record, foram realizados em um teatro com a presença do público. A forma do teatro, mais do que a tradição do rádio, caracterizava a cena musical televisionada. Os gestos amplos, que lhe renderam o apelido de Eliscóptero, eram midiatisados pela televisão em planos abertos (imagem-ação e imagem-percepção) e as expressões faciais eram superdimensionadas pelos planos fechados no rosto da cantora (imagem-afecção). Assim, respectivamente, a articulação entre imagem-ação, percepção e imagem-afecção⁶ constituía o propriamente televisivo. Mas a forma da cena respondia preponderantemente às gramáticas dos musicais realizados em teatros.

⁵ Convém referir que a relação, aqui, não é de causalidade, mas de abertura de possibilidades intensivas virtuais. Defende-se a tese de que Elis Regina na televisão expressa devires de imagem-música, mas não é sua causa.

⁶ Os conceitos de imagem-ação, imagem-afecção e imagem-percepção são desenvolvidos por Gilles Deleuze em *Imagem-Movimento* (1984).

Essa articulação entre cena teatral e imagem televisiva produziu em Elis o que aqui se chama de voz-corpo-televisivo⁷. Desde uma perspectiva televisiva, o que importam são as formas audiovisuais; desde uma perspectiva teatral, as inter-relações formais entre corpo e movimento em um espaço circunscrito. Entretanto, chama-se, aqui, a atenção para o aspecto das influências recíprocas entre essas diferentes formas. Ainda que a forma televisiva seja preponderante, podem-se observar – e Augusto de Campos já o fez em seu famoso ensaio *O balanço da Bossa e outras bossas* (1993) – como se foram produzindo modificações no modo como Elis cantava a partir de sua experiência com a televisão e, agora, para além do texto de Campos, como se foram produzindo modificações no modo como a televisão criava gêneros e formatos musicais recriando a própria música até que tal procedimento se autonomizou em imagem-música, com a invenção de programas computacionais.

Pode-se, então, agora, retomar os níveis de conexão entre imagem e música para dizer que o terceiro consiste na autonomização dos modelos que simulam, a partir de programas computacionais, textualizações que se apresentam diferentemente aos sentidos: como som (musical) e como imagem (televisiva). A análise dessas imagens-músicas não se constitui em objetivo deste artigo cujo foco recai sobre os mecanismos semióticos que as tornaram possíveis, que as ensaiaram previamente. Entretanto, pode-se levantar a hipótese de que, assim como os Disk Jockeys (DJs) têm recriado antigas canções⁸ conferindo-lhes forma digital, para o quê transformam as gravações originais em matérias para nova formalização, também parece pertinente imaginar a possibilidade de se inventarem audiovisualidades digitais a partir da matéria audiovisual disponível⁹, trazendo de novo à cena viva, como matéria, o que parecia estar definitivamente calado.

Essas três formas de conexão – relativas ao meio, às inter-relações entre linguagens e às simulações – oferecem o mapa onde se produzem os devires de audiovisualidades que se quer estudar e evidenciam os processos semióticos como recriação formal e infinita da matéria.

Referências bibliográficas

1. ALMEIDA, Júlia. *Estudos deleuzeanos da linguagem*. Campinas: Ed Unicamp, 2003.
2. BECKER, Howard S. Mundos artísticos e tipos sociais. In: VELHO, Gilberto. *Arte e sociedade: ensaios de sociologia da arte*. Rio de Janeiro: Zahar, 1977.
3. CAMPOS, Augusto de. *Balanço da bossa e outras bossas*. São Paulo: Perspectiva, 1993.
4. DELEUZE, Gilles. *A Imagem-Movimento*. São Paulo: Brasiliense, 1985.
5. DELEUZE, Gilles. *A Imagem-Tempo*. São Paulo: Brasiliense, 1985.
6. DELEUZE, Gilles. *Crítica e clínica*. São Paulo: Ed. 34, 1997.
7. DELEUZE, Gilles. *Périclès et Verd: la philosophie de François Châtelet*. Paris: Minuit, 1988.
8. DERRIDA, Jacques. *Ecografías de la televisión*. Buenos Aires: Editora Universitária, 1998.
9. GUATTARI, Félix. *Cartographies schizoanalytiques*. Paris: Éditions Galilée, 1989.
10. HJELMSLEV, Louis. *Prolegômenos a uma teoria da linguagem*. São Paulo: Perspectiva, 1961.
11. JOST, François. *Seis lições sobre a televisão*. Porto Alegre: Sulina, 2004.
12. KILPP, Suzana. *Ethicidades televisivas: sentidos identitários na TV*. São Leopoldo: Unisinos, 2004.
13. SANTAELLA, Lúcia; NÖTH, Winfried. *Imagem: cognição semiótica, mídia*. São Paulo: Iluminuras, 1998.
14. SILVA, Alexandre Rocha da. *Elementos para uma comunicação pós-midiática*. (tese). Unisinos, 2003.
15. SILVA, Alexandre Rocha da e PELLEZ, Vinícius da Silva. Memória audiovisual brasileira. In: DUARTE, Elizabeth Bastos e CASTRO, Maria Lília Dias de. *Televisão entre o mercado e a academia*. Porto Alegre: Sulina, 2006.

⁷ Aqui não se está negando a influência do rádio – sobretudo das grandes cantoras da Rádio Nacional, como Ângela Maria – sobre a elaboração do canto de Elis Regina. Está-se, apenas, priorizando os traços de influência próprios da televisão.

⁸ Há gravações recriadas por DJs a partir de canções de Elis Regina, Tom Jobim, Billie Holiday, entre outros.

⁹ O acesso à matéria audiovisual, no Brasil, é dificultado pelo seu estatuto predominantemente privado. Tal problema, político, embora difícil, não inviabiliza, teoricamente, a produção de imagens-músicas. Outras reflexões sobre tais problemas políticos podem ser encontradas no artigo *Memória Audiovisual Brasileira* (2006)