## CRIAÇÃO MUSICAL E TECNOLOGIAS:

teoria e prática interdisciplinar

SÉRIE PESQUISA EM MÚSICA NO BRASIL

volume 2

## **ANPPOM**

damián keller (org)
rogério budasz (org)
josé fornari
victor lazzarini
marcelo m. wanderley

daniel luís barreiro jônatas manzolli adolfo maia jr rodolfo coelho de souza álvaro barbosa



ISBN 978-85-63046-01-7

# CRIAÇÃO MUSICAL E TECNOLOGIAS: teoria e prática interdisciplinar

## SÉRIE PESQUISA EM MÚSICA NO BRASIL VOLUME 2

## **ANPPOM**

damián keller (org.)

rogério budasz (org.)

jonatas manzolli

josé fornari

victor lazzarini

marcelo wanderley

daniel luís barreiro

jonatas manzolli

adolfo maia jr.

rodolfo coelho de souza

álvaro barbosa



## ASSOCIAÇÃO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM MÚSICA

### Diretoria 2009-2011

Presidente: Sonia Ray (UFG) Ia Secretária: Lia Tomás (UNESP) 2a Secretária: Cláudia Zanini (UFPR) Tesoureira: Sonia Albano de Lima (FCG)

### **Conselho Fiscal**

Denise Garcia (UNICAMP) Martha Ulhôa (UNIRIO) Ricardo Freire (UnB) Acácio Piedade (UDESC) Jonatas Manzolli (UNICAMP) Fausto Borém (UFMG)

#### Conselho Editorial

Rogério Budasz (UCR) Paulo Castagna (UNESP) Norton Dudeque (UFPR) Acácio Piedade (UDESC)

# CRIAÇÃO MUSICAL E TECNOLOGIAS: teoria e prática interdisciplinar

SÉRIE PESQUISA EM MÚSICA NO BRASIL VOLUME 2

**ANPPOM** 

## CRIAÇÃO MUSICAL E TECNOLOGIAS:

## teoria e prática interdisciplinar

#### CAPA

Detalhe da escultura sonora "Green Canopy". Damián Keller e Patricia Tinajero

#### CONSELHO EDITORIAL:

Álvaro Barbosa, Damián Keller, Daniel Luís Barreiro, Eduardo Reck Miranda, Jônatas Manzolli, José Fornari (Tuti), Marcelo M. Wanderley, Marcelo S. Pimenta, Mikhail Malt, Rodolfo Coelho de Souza, Victor Lazzarini

Criação musical e tecnologias: teoria e prática interdisciplinar / Damián Keller (prefácio e organização), Rogério Budasz (organização e editoração) – Goiânia : ANPPOM, 2010.

202 p. : il. : 21 cm. ISBN 978-85-63046-01-7

I. Música. 2. Musicologia. 3. Composição (Música). 4. Música – Instrução e Ensino. 5. Música – Interpretação. I. Keller, Damián. II. Budasz, Rogério. III. Fornari, José. IV. Lazzarini, Victor. V. Wanderley, Marcelo M. VI. Barreiro, Daniel Luís. VII. Manzolli, Jônatas. VIII. Maia Jr, Adolfo. IX. Souza, Rodolfo Coelho de. X. Barbosa, Álvaro. XI. Título.

CDD 781

ANPPOM
Associação Nacional de Pesquisa e
Pós-Graduação em Música

www.anppom.com

Printed in Brazil 2010

Prefácio	1
	Damián Keller
Percepção, cognição e afeto musical	6
	José Fornari (Tuti)
Técnicas fundamentais para instrumentos computacionais de síntese e processamento de áudi	39 <b>o</b>
	Victor Lazzarini
Projeto e utilização de instrumentos musicais digita	<b>is</b> 70
Mar	celo M. Wanderley
Composição com modelos sonoros: fundamentos e aplicações eletroacústicas	97
Do	nniel Luís Barreiro, Damián Keller
Composição textural algorítmica via sistemas dinâm	nicos 127
	Jônatas Manzolli, Adolfo Maia Jr.
Da interação entre sons instrumentais e eletrônicos	149
Rodol	fo Coelho de Souza
Performance musical em rede	180
	Álvaro Barbosa
Glossário	201



## Criação musical e tecnologias: o que há e o que virá

DAMIÁN KELLER

Caro mestrando e doutorando,

A vida do compositor-pesquisador é dura. Você estuda dez anos de violino, piano, agogô. Aprende notação musical, escalas, arpejos. Bate palmas e faz "pa, pa, pa" até ficar rouco e com calos nas mãos. Aí você chega na universidade, no seu primeiro dia de aula de composição, pensando "agora é moleza, já sei tudo que tenho que saber de música". Surpresa, surpresa! O que você aprendeu ainda não é suficiente para criar teu primeiro som. E agora, o que fazer? Você precisa entender a anatomia do sistema perceptivo (Capítulo 1), as diversas relações entre matemática e música (Capítulo 5), a aplicação de conceitos como gesto, textura e espectro-morfologia na organização musical (Capítulo 6), e ter ao menos um conhecimento básico de linguagens de programação (Capítulo 2). Enfim, se você já passou por isso, este livro vai servir para suprir algumas dessas carências.

Mas o objetivo deste volume não é somente compensar a falta de textos de consulta sobre computação musical, criação sonora e técnicas composicionais. Ao encarar a pesquisa bibliográfica para sua monografia, você vai descobrir (ou já está descobrindo) que não há consenso nos termos, métodos e nem nos fundamentos epistemológicos da nossa área. A atividade composicional com tecnologias envolve perspectivas múltiplas, incluindo os métodos herdados de um milênio de escrita musical, mas também os últimos avanços científicos e tecnológicos em áreas que aparentemente não têm relação direta com a música. Por isso tivemos o cuidado de acrescentar uma lista de termos específicos com suas respectivas definições (o Glossário). Essas definições são relativas ao contexto do capítulo para a qual foram escritas, e passaram pelo crivo de um painel de especialistas (o nosso Conselho Editorial). Ou seja, são conceitos que vêm sendo discutidos no trabalho de vários pesquisadores e que eventualmente podem se tornar uma referência útil para desenvolvimentos futuros.

E falando em referências, todos os capítulos incluem bibliografia abrangente, desde textos fundacionais de autores como Schafer (2001), Varèse (1959) e Xenakis (1971) – até textos técnicos – de Dodge e Jerse (1985), Roads (1996) e outros. Muitos artigos recentes em português,

publicados em eventos como o Congresso Anual da ANPPOM ou o Simpósio Brasileiro de Computação Musical (SBCM), também foram discutidos. Para os casos em que não foi possível abordar um determinado trabalho no corpo do capítulo, separamos um espaço de Referências Complementares. Sem dúvida, muitos trabalhos importantes não foram abordados, mas com certeza haverá espaço para mais contribuições nos próximos volumes da série.

Escrevendo estas linhas me confronto com um problema que encontrei na época em que eu dava meus primeiros passos na área de computação musical. Nessa época não existia internet, os sistemas para fazer música em tempo real (hoje diríamos sistemas síncronos) eram poucos e a maior parte da pesquisa estava concentrada em alguns estúdios vinculados às grandes universidades. Ou seja, era necessário gostar muito de tecnologia para se animar a enfrentar a montanha de dificuldades que você encontrava na hora de fazer música por computador. Nessa situação, a pergunta era: por onde começar?

O contexto mudou quantitativa e qualitativamente. Dispositivos computacionais formam parte do dia-a-dia da maioria da população (o telefone celular é o exemplo mais citado, mas outros tipos de dispositivos estão sendo incorporados a uma variedade de objetos e artefatos de uso cotidiano). A cobertura de redes sem fio com acesso gratuito está em rápida expansão, prevendo-se a disponibilidade em todos os estados do Brasil. Hoje praticamente todos os alunos da rede escolar brasileira têm contato cotidiano com computador e internet. E é ali, na rede global, onde você pode encontrar informações atualizadas sobre praticamente todas as áreas da computação musical. Com o acesso maciço aos meios tecnológicos por parte de grupos sociais diversos, e tendo a possibilidade de ampliar o fazer musical a contextos e situações que há poucos anos formavam parte da ficção científica, volta a surgir a pergunta: por onde começar?

Uma leitura rápida dos anais recentes dos principais simpósios de música e computação musical no Brasil - incluindo o Congresso da ANPPOM e o SBCM – mostra que mesmo em trabalhos interessantes, existe uma tendência para o isolamento da pesquisa em música em relação às outras áreas do conhecimento. Esse isolamento é particularmente negativo para três aspectos do estudo do fazer criativo-musical: a percepção, a socialização e o desenvolvimento de ferramentas (engenharia musical). A relevância dos processos perceptivos na criação musical é amplamente discutida na literatura, configurando uma área independente de pesquisa: a psicologia da música (também conhecida como cognição musical) (FORNARI, 2010). O objetivo dessa área é entender como o nosso sistema auditivo (incluindo os processos do sistema periférico – o ouvido – e os processos do sistema nervoso central) processa estímulos sonoros que eventualmente são classificados como música.

Os aspectos sociais do fazer musical têm sido amplamente estudados pela Etnomusicologia ou Antropologia Musical (ARAÚJO *et al.*, 2008; REYNOSO, 2007). Existem, inclusive, estudos etnográficos específicos focalizando aspectos da comunidade musical tecnológica (BORN, 1995). No entanto, esse conhecimento geralmente não encontra aplicabilidade na criação musical. É possível que com a atenção crescente dada ao contexto no fazer musical (KELLER e CAPASSO, 2006; SCHAFER, 2001; TRUAX, 2002), a contribuição da Antropologia Musical comece a ser melhor aproveitada nos processos criativos.

O terceiro aspecto ausente nas discussões acadêmicas sobre criação musical é a questão prática - mas perigosamente limitadora - da implementação de técnicas composicionais. O trabalho criativo-musical geralmente é pensado como aplicação de técnicas já existentes. Essas técnicas podem ser combinadas de forma livre, podem ser "subvertidas" mudando a sua funcionalidade original, ou podem ser aplicadas a contextos musicais previamente impensados, mas raramente o compositor desenvolve o seu próprio instrumento de trabalho. Com o surgimento de ambientes livres para programação musical (Csound, Pure Data, Nyquist e outros) e com o barateamento dos sistemas de controle (teclados e superfícies com suporte para o protocolo MIDI, e placas lógicas programáveis), vários pesquisadores têm começado a desenvolver suas próprias ferramentas (ver Capítulos 2 e 3). É justamente nessa atividade – que podemos chamar de engenharia musical – onde residem os grandes perigos e as grandes oportunidades. Os perigos consistem, por um lado, na falta de aplicabilidade desse trabalho por conta da fragmentação e multiplicidade dos métodos utilizados, com resultados restritos a uma obra ou a um grupo pequeno de obras. Por outro lado, a adoção de soluções prontas (os enlatados comerciais) impõem limitações na atividade criativa, com a consequente homogeneização do resultado musical. O caso mais patético é a música feita através de sequenciadores ou editores de notação utilizando o padrão General MIDI para renderização. Esse tipo de enlatados determina um pensamento musical condicionado representação métrico-tonal e reduz as possibilidades de controle fino de nuances interpretativas fornecidas pela escrita instrumental. Em oposição a essa tendência, os trabalhos mostrados nos capítulos deste volume fornecem bons exemplos de como evitar essa armadilha. Desde uma perspectiva interdisciplinar integrada, o desenvolvimento de novas ferramentas fomenta a formulação de hipóteses que, ao serem testadas, criam novas oportunidades para compositores, intérpretes e público.

Neste volume, você encontrará material para embasar o desenvolvimento e a aplicação de tecnologia na pesquisa composicional. No capítulo 1, Tuti Fornari discute os aspectos perceptuais, cognitivos e afetivos do fazer musical. Os aspectos perceptivos e cognitivos estão relacionados a como a música é percebida. Os aspectos afetivos são estudados através da observação do efeito emocional nos ouvintes. No capítulo 2, Victor Lazzarini

mostra técnicas e implementações de instrumentos de síntese e processamento de dados musicais, focalizando a manipulação sonora, No capítulo 3, Marcelo Wanderley coloca a questão de como aproveitar musicalmente a geração sonora em tempo real através do desenvolvimento de dispositivos de controle com capacidade para captar os gestos dos intérpretes. Daniel Barreiro e Damián Keller tratam da fundamentação e aplicação de técnicas composicionais eletroacústicas, mostrando exemplos de algoritmos e resultados sonoros extraídos das suas obras. Jônatas Manzolli e Adolfo Maia Jr. apresentam dois modelos de composição de texturas sonoras. O primeiro modelo é controlado por uma sequência numérica gerada iterativamente. O segundo modelo consiste em uma matriz de parâmetros para controlar as camadas do processo textural. No capítulo 6, utilizando diversos exemplos composicionais, Rodolfo Coelho de Souza aborda tópicos relativos à interação entre sons instrumentais e eletrônicos, incluindo conceitos como tipologia espectral, tipologia morfológica, movimento virtual, sintaxe, parataxe, e contraste. No capítulo 7, Álvaro Barbosa trata da latência de comunicação (através da conexão em rede ou provocada por computação intensa) e dos seus efeitos na sincronização dos dados musicais. Com base numa concepção instrumental, a performance colaborativa requer comunicação em tempo real entre executantes. Através de uma série de experimentos, Barbosa mostra que - com prática - os músicos tendem a melhorar a sua capacidade para tocar sob condições adversas, sugerindo que existe uma relação inversa entre tempo musical fixo e tolerância a latência.

Resumindo, o volume abrange questões cognitivas, composicionais e tecnológicas com resultados relevantes para várias subáreas da pesquisa em criação musical com tecnologia. A partir do estudo detalhado dos métodos e exemplos de cada capítulo, você pode abordar aspectos da implementação de sistemas síncronos e assíncronos, aplicar métodos composicionais e avaliar a relevância das técnicas e dos enfoques atuais para a pesquisa sonoro-musical.

Esperamos que os conceitos e os sons deste livro abram novos caminhos na sua prática musical. Se a leitura deste texto lhe der ideias para fazer ao menos um barulhinho interessante, teremos atingido o nosso objetivo.

Um abraço,

Damián

#### Referências

ARAÚJO, Samuel; PAZ, Gaspar; CAMBRIA, Vicenzo (orgs.) *Música em Debate: Perspectivas Interdisciplinares.* Rio de Janeiro: Mauad X - FAPERJ, 2008.

DODGE, Charles; JERSE, Thomas. Computer music. New York: Schirmer, 1985.

FORNARI, José. Percepção, cognição e afeto musical. In: KELLER, Damián; BUDASZ, Rogério (orgs.). *Criação musical e tecnologias: teoria e prática interdisciplinar*, p. 6-38, Goiânia: ANPPOM, 2010.

KELLER, Damián; CAPASSO, Ariadna. New concepts and techniques in ecocomposition. *Organised Sound*, v. 11 (2006), p. 55-62.

LAZZARINI, Victor. Técnicas fundamentais para instrumentos computacionais de síntese e processamento de áudio. In: KELLER, Damián; BUDASZ, Rogério (org.). *Criação musical e tecnologias: teoria e prática interdisciplinar*, p. 39-69, Goiânia: ANPPOM, 2010.

REYNOSO, Carlos. *Antropología de la Música: De los Géneros Tribales a la Globalización. Volumen 1: Teorías de la Simplicidad.* Buenos Aires: Editorial SB, 2007.

ROADS, Curtis. *The Computer Music Tutorial*, Cambridge, MA: MIT Press, 1996.

SCHAFER, R. Murray. *A afinação do mundo.* São Paulo: Ed. UNESP, 2001.

TRUAX, Barry. Genres and techniques of soundscape composition as developed at Simon Fraser University. *Organised Sound*, v. 7, n. 1 (2002), p. 5-14.

VARÈSE, Edgar. *Rhythm, form and content,* 1959. [Palestra - excertos publicados em *The Liberation of Sound*]. In: CHILDS, Barney; SCHWARTZ, Elliott (orgs.) *Contemporary composers on contemporary music,* p. 195-208. Nova York: Holt, Rinehart, and Winston, 1967.

XENAKIS, Iannis. *Formalized music: thought and mathematics in composition.* Bloomington: Indiana University Press, 1971.

## Percepção, cognição e afeto musical

JOSÉ FORNARI (TUTI)

e você faz parte do grupo de indivíduos com audição normal (considerado acima de 90% da média da população) então, durante toda sua vida de ouvinte, estará sendo incessantemente exposto à informação sonora. Estudos sugerem que desde aproximadamente as 18 semanas de gestação (NARA, 1996, p. 35), o feto humano já é sensível a sons, tanto os internamente gerados pelo organismo da mãe quanto por sons externos, como ruídos, músicas e vozes de outras pessoas. Se permanecermos por um certo período tempo num ambiente extremamente silencioso, como o interior de uma câmara anecóica, passamos a ouvir, muitas vezes para nosso espanto, a imensa quantidade de ruídos gerados pelo nosso próprio corpo, como a movimentação de músculos, o batimento cardíaco, a respiração, a circulação sanguínea e um ruído aproximadamente constante gerado pelo próprio ouvido, que, se ocorre em grau acentuado é uma anomalia conhecida pelo nome de tinnitus.1 Ouvimos sons estando ou não conscientes de sua existência, sendo que, se a nossa atenção está voltada ao estímulo sonoro, dizemos que não estamos apenas ouvindo mas também "escutando". Podemos também, e com frequência, escutar sons que não provêm de fontes sonoras físicas, tais como os sons de músicas e vozes que lembramos ou que sonhamos. Podemos ouvir sem escutar, escutar sem compreender, ou compreender sem nos emocionar.

¹ Zumbido, também chamado de acúfeno ou tinido, que é uma sensação auditiva cuja fonte não advém de estímulo externo ao organismo, é um sintoma muitas vezes associado a várias formas de perda auditiva.

Isto nos faz pensar sobre a natureza do som. Onde está o som que escutamos? Está fora de nós, no mundo externo, sob a forma de ondas de variação de pressão do meio propagante (o ar) que nos atingem; compressões e expansões longitudinais do meio elástico (o ar), aproximadamente periódicas que ocorrem dentro de um específico limite de intensidade, frequência e tempo? Ou será que o som está dentro de nós, no universo subjetivo de nossas mentes, sob a forma da percepção, interpretação e reação às tais oscilações de pressão acústica; no processo fisiológico que nos permite reconhecer, entender, lembrar e até nos emocionar com aquilo que ouvimos? Consideramos o som como a informação que permeia o domínio externo (acústico) e o interno (subjetivo). O som é a informação que percebemos pelo sentido da compreendemos pelos processos cognitivos eventualmente pode nos causar reações afetivas, através da evocação de emoções. Dentro desta perspectiva, o som é informação multidimensional seus e (unidimensionais) organizam-se em uma das três categorias, aqui referidas por percepção, cognição e afeto.

Sons podem comunicar conceitos dentro dessas três categorias. Em termos de percepção, estamos frequentemente dependendo de sons para orientar nossas ações, nos preparar situações inesperadas para e assim tentar corretamente eventos futuros. É um dos processos mentais que realizamos ao atravessar uma avenida movimentada ou ao ouvir as trovoadas de uma tempestade se aproximando. No campo da cognição sonora, temos o exemplo das linguagens humanas: а capacidade de comunicação de conceitos complexos, muitas vezes intangíveis e abstratos, considerada por diversos filósofos como um atributo que a distingue da comunicação das demais espécies animais e assim viabiliza a formação de culturas e também das artes, como é o caso da música. Esta é a arte que organiza eventos sonoros ao longo do tempo a fim de compor um discurso estético. A música apresenta três perspectivas: a composição, a performance e a análise. A composição tradicionalmente lida com o processo de concepção e implementação atemporal de uma estrutura 8

musical.<sup>2</sup> A performance trata da interpretação musical dessa estrutura, através de sua manifestação ao longo do tempo. A análise estuda os elementos da estrutura composicional e seu encadeamento temporal durante a performance expressiva.<sup>3</sup>

A música pode ser estudada através dos seus aspectos independentes e dependentes do contexto musical; ou seja, da informação prévia (que o ouvinte já possui) sobre o material musical ouvido, bem como de sua expectativa sobre aquilo que será escutado na música. Essas são descritas como grandezas unidimensionais, escalares ou funções do tempo, que expressam (a variação de) uma característica particular e independente da música (ou de um trecho musical). Esses são chamados de aspectos musicais. Tais aspectos são separados em dois grupos: independentes e dependentes do contexto musical; ou seja, da informação prévia (que o ouvinte já possui) sobre o material musical ouvido, ou de sua expectativa sobre aquilo que será escutado na música. O contexto musical está assim relacionado aos processos de memorização da mente humana e a consequente distinção e disposição de eventos na linha do tempo, estendendo-se entre passado, presente e futuro. A memória lida com o entendimento dos eventos musicais ocorridos no tempo presente da música, e sua comparação com os eventos que ocorreram no passado. Ao escutarmos uma música, existe uma certa duração do tempo do um intervalo ou janela temporal, que discurso. espontaneamente interpretamos como o "agora", o momento atual da escuta musical, onde todos os eventos que ocorrem dentro deste intervalo são interpretados pela mente como estando ocorrendo no momento presente. Durante a escuta musical, podemos lembrar de alguns aspectos da informação passada bem como criar predições ou expectativas de uma

<sup>&</sup>lt;sup>2</sup> Para uma discussão dos processos composicionais contemporâneos, ver conceitos de "no-tempo" e "fora-do-tempo" em Barreiro e Keller (2010)

<sup>&</sup>lt;sup>3</sup> Utilizamos aqui o termo "expressivo" num contexto tradicional. Existem correntes composicionais que não focalizam na expressão musical, por exemplo, John Cage (4' 33", 1952) e o minimalismo de Morton Feldman, entre outros.

possível sucessão de eventos correspondentes no futuro. Além disso, mesmo antes de começarmos a escutar uma música, já temos uma série de expectativas referentes a aspectos como o gênero musical, o estilo da performance, e a qualidade sonora (gravação, acústica da sala, etc.). Grosso modo, podemos dizer que a memória de curta-duração está relacionada à expectativa musical intrínseca, que ocorre durante a escuta, enquanto que a memória de longa-duração, a uma expectativa extrínseca, iniciada anteriormente à escuta, com influência sócio-cultural e relacionada ao gênero musical.

A música é tradicionalmente<sup>4</sup> uma arte de eventos que se desenrolam ao longo do tempo, o que a distingue de outras formas artísticas, como a pintura ou a escultura, onde a informação artística encontra-se disposta atemporalmente, condensada no objeto artístico, onde um observador tem a total liberdade de examinar a obra em seu próprio "ritmo", determinando a sequência de passos na apreciação dos detalhes desta obra. Já a música é estritamente atrelada à orientação temporal de seu discurso, do qual o contexto musical totalmente dependente. **Podemos** discrepantes versões de uma peça musical conhecida e ainda assim facilmente a identificarmos, mesmo alterando aspectos intensidade. tonalidade. andamento. (orquestração) ou métrica (performance). No entanto, se invertermos o sentido de sua performance (se executarmos uma partitura de trás para frente; do fim para o início) a obra se tornará irreconhecível. Esta será escutada como uma música totalmente diferente de sua versão original. Pode-se constatar que o contexto musical é mais dependente da ordenação temporal que as artes visuais no tempo, como o cinema, onde, se assistirmos um trecho de trás para frente, ainda assim teremos uma noção do seu contexto original. Talvez por ter o contexto tão atrelado a uma ordenação temporal, a música seja tão eficiente na criação e manipulação de expectativas. Estas

<sup>&</sup>lt;sup>4</sup> Nota-se, todavia, que, na música contemporânea, existem muitos exemplos de manipulações composicionais que trabalham com inversão (retrogradação) de eventos musicais, tais como o serialismo e música concreta.

são associadas à evocação de emoções que, numa abordagem evolutiva, cumprem o papel de mover o indivíduo à ação, de modo a garantir seu bem-estar e até sua sobrevivência. De um modo geral, as emoções evocadas por expectativas estão associadas à predição de eventos futuros. Na psicologia das expectativas, emoções são consideradas como estratégias evolutivas de incentivo a modelos mentais que incorram em predições corretas e dirimam modelos que resultem em predições incorretas de eventos. A composição musical viabiliza a manipulação eficiente de expectativas contextuais intrínsecas à sua prosódia, e é também embasada num contexto maior, extrínseco à obra, seu gênero musical. Isto permite que a música seja utilizada como meio de criação de um discurso de emoções, o que vem de longa data, sendo utilizado por outros meios de comunicação e artes (canções, trilhas sonoras, campanhas publicitárias, etc.) como forma de intensificar uma mensagem afetiva.

10

Neste capítulo iremos discutir os aspectos perceptuais, cognitivos e afetivos da música; como estas categorias são definidas, a distinção entre seus aspectos e seus respectivos papéis no processo de criação, expressão e apreciação musical. Aspectos perceptivos são aqueles que descrevem a maneira como a informação sonora é captada como estímulo acústico (externo) pelo sistema binaural (formado pelo par de ouvidos). Tais aspectos são estudados pela área da ciência conhecida como psicoacústica. Características psicoacústicas ocorrem em intervalos de tempo suficientemente pequenos, anteriores à formação de um modelo de memorização de informação sonora (assim, não existe distinção entre aspectos psicoacústicos sonoros e musicais). Tais aspectos estão associados a um intervalo de tempo conhecido na psicoacústica como intervalo da persistência auditiva, considerado em torno de 0,1s de duração (JACOBS, 1972). Eventos sonoros distintos, separados por intervalo de tempo menor que o da persistência auditiva percebidos são como um único evento Correspondentemente, o limite mínimo de percepção tonal (ou seja, da frequência de um evento sonoro aproximadamente periódico) é cerca de 20 Hz de frequência, o que equivale a um período de tempo de 0,05s, enquanto que a persistência auditiva é cerca de aproximadamente 0,1s. Pode-se assim supor

11

que a percepção tonal seja uma estratégia evolutiva para dar continuidade à percepção da variação do intervalo de tempo entre eventos sonoros. Isto implica dizer que a predição de eventos sonoros, de uma mesma fonte sonora, distanciados no tempo por intervalos maiores que o da persistência auditiva, pode apresentar o contexto de ritmo, enquanto que eventos sonoros separados por intervalos menores que o da persistência auditiva apresentam apenas o contexto de altura, tom ou *pitch*. A fronteira entre o domínio contextual rítmico e tonal encontra-se aproximadamente estendida entre 0,05 e 0.1s.

Aspectos cognitivos tratam dos processos sonoros internos, ou mentais. São aqueles formados por eventos sonoros que ocorrem acima de um intervalo de tempo suficiente para se estabelecer uma memória de sua ocorrência e assim, na música, passam a ter significado ou contexto musical. No processo cognitivo, existe um intervalo ou janela de tempo que corresponde à noção do agora musical. Esta tem suas origens nos estudos do psicólogo William James, que desenvolveu o conceito de "presente especial" (JAMES, 2007). Segundo James, este intervalo varia de indivíduo para indivíduo, bem como pela modalidade da informação sonora, se esta é ambiental, linguística ou musical, mas é considerada como tendo cerca de 1 a 3 segundos de duração, o que é relacionado à memória de curta-duração (POIDEVIN, 2000).

Aspectos afetivos são aqueles que lidam com a evocação de emoções pela música através do discurso de expectativas. Existe uma distinção entre esses e os aspectos cognitivos que descrevem emoções. Aspectos afetivos lidam com emoções de fato evocadas enquanto que os aspectos cognitivos estão necessariamente associados às emoções constatadas mas não evocadas na mente do ouvinte. Por exemplo, podemos escutar uma animada trilha sonora para um espetáculo de circo e constatar que se trata de uma música de caráter alegre, mas sem experimentarmos a emoção de alegria, bem como escutarmos um *Requiem* sem nos sentirmos tristes ou deprimidos. Aspectos afetivos agem em intervalos de tempo maiores, onde se tem a influência da memória de longaduração, mais especificamente, da memória declarativa e procedural; que nos permite voluntariamente lembrar de fatos

e passagens, objetos e movimentos. Esta é relacionada à evocação de emoções através da escuta e, na música, com o reconhecimento do estilo da performance do interprete ou com o gênero da música.

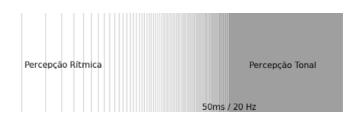
## A percepção musical

Ao passarmos em frente a uma construção, escutamos sons que não necessariamente compreendemos. Não sabemos exatamente qual é a sua origem nem associamos máquinas ou processos específicos que poderiam gerá-los. No entanto, percebemos claramente suas intensidades sonoras, detectamos se tais sons apresentam uma altura (pitch), se são sons "abafados", se "ásperos" ou existe reverberação, Percebemos uma série de aspectos destes sons, sem necessariamente os compreender. Esses são os aspectos perceptuais sonoros. Na música, a percepção sonora diz respeito à audição dos aspectos que descrevem características psicoacústicas do material escutado. Esses são também chamados de aspectos de baixo-nível, ou não-contextuais, em contraste aos aspectos de alto-nível ou contextuais, tais como os cognitivos e afetivos. O termo "baixo-nível" se refere às características intrínsecas do sinal acústico que ocorrem na escala do limiar da persistência auditiva e definem a percepção da informação sonora pela audição. Desse modo, não existe uma distinção entre percepção sonora e musical, uma vez que estes ocorrem numa escala temporal pequena, anterior à formação de memória. São chamados de descritores de baixo-(Low-Level Descriptors), nível. 011 LLD algoritmos computacionais capazes de calcular aspectos psicoacústicos similar ao processamento fisiológico da audição humana. Estes são utilizados pela indústria, como no padrão MPEG-7, que utiliza 17 LLDs para a descrição de áudio (MANJUNATH et al., 2002). Este é composto por descritores no domínio do tempo e da frequência, tais como intensidade, ataque, densidade espectral, frequência fundamental (altura), etc.

A percepção musical inicia-se pela audição humana, que trata da captação simultânea pelos dois ouvidos do sinal acústico; as sucessivas ondas de oscilação aproximadamente periódica de compressão e expansão do meio elástico, que é

12

normalmente o ar (ainda que ouçamos em outros meios materiais, tais como embaixo da água, através de objetos sólidos, como paredes, etc.). Este primeiro estágio da audição traduz a informação de oscilação mecânica em sinais elétricos, através de disparos de potencial neurológico que caminham pelo nervo auditivo (o nervo vestíbulo-coclear) ao cérebro. Este processo ocorre continuamente, em tempo real, e se estende no domínio do tempo desde a percepção de eventos sonoros rítmicos, onde estes se encontram espaçados por intervalos acima da persistência auditiva, até intervalos menores que o período equivalente ao início da percepção tonal (cerca de 50ms ou 20Hz). Esta região, mostrada na Figura 1, demarca a fronteira da percepção rítmica para tonal.



**Fig. 1:** Fronteira da percepção temporal sonora. Em torno de 50ms de período, ou 20 Hz de frequência, a percepção do intervalo de tempo entre eventos sonoros passa de rítmica (domínio do tempo) para tonal (domínio da frequência).

Em linhas gerais, a audição traduz a informação sonora de sinais acústicos (variações de pressão), para mecânicos (variações de deslocamentos) e, por fim, em sinais elétricos que são transmitidos pelo nervo auditivo ao cérebro. Os sinais acústicos são variações aproximadamente periódicas de pressão atmosférica (o meio elástico onde o som normalmente se propaga) que são captadas pela orelha, trafegam pelo canal auditivo, são transformadas em vibrações mecânicas no ouvido médio e finalmente em sinais elétricos, no ouvido interno. O sentido da audição é, algumas vezes também referenciado como sistema binaural pois, de fato, este utiliza a informação do par de ouvidos para também captar a localização espacial das fontes sonoras. Cada ouvido é um sistema independente,

14

composto por três partes - ouvidos externo, médio e interno. O ouvido externo é composto pelo pavilhão (orelha) e o conduto auditivo. Além de proteger as camadas internas do ouvido, este também apresenta a propriedade de filtrar o som de modo a realçar as frequências mais importantes para a comunicação humana e ajudar na localização da posição da fonte sonora no espaço. O ouvido médio é composto pelo tímpano, uma membrana que vibra com as ondas acústicas do som e transforma as oscilações de pressão do ar em vibrações mecânicas. O tímpano está conectado a um conjunto de minúsculos ossos (o martelo, a bigorna e o estribo) conectados a músculos que os acomodam, ajudando a atenuar ou amplificar sua vibração mecânica e a transportá-la para o ouvido interno, através de uma abertura chamada de janela oval. O ouvido interno é composto pela cóclea, um órgão espiralado, e pelo sistema vestibular, formado por canais semicirculares responsáveis pela sensação de equilibro e deslocamento corporal. A cóclea é responsável pela tradução das vibrações mecânicas vindas do ouvido médio (através da conexão do estribo com a janela oval) em sinais elétricos. Dentro da cóclea encontra-se o órgão de Corti, que, se desenrolado, possui um formato aproximadamente de cunha. Este contém de 15.000 a 20.000 células ciliadas que produzem potencial elétrico com a movimentação mecânica de seus cílios (os estereocíclios). Estes cílios estão mecanicamente conectados a uma membrana fixa, a membrana tectória. O som faz com que as regiões do órgão de Corti entrem em vibração. De acordo com as componentes presentes no som, diferentes regiões entrarão em ressonância. As células ciliadas, conectadas à região que vibra, respondem gerando impulsos elétricos que são transportados pelo nervo auditivo ao tronco encefálico, que inicia o processo de interpretação desses sinais elétricos em som. A Figura 2 mostra uma visão simplificada do processo de captação da informação sonora pelo ouvido, onde esta é traduzida de informação acústica (vibrações de pressão atmosférica) em mecânica, pelo ouvido médio, e elétrica, pelo ouvido interno.

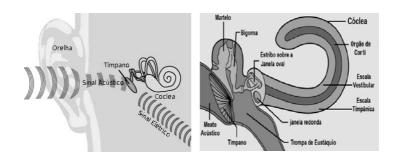


Fig. 2: A captação da informação sonora pelo ouvido humano. A figura à esquerda mostra o sinal acústico, composto pelas ondas longitudinais de variação de pressão do ar, entrando pela orelha, sendo transformadas em vibração mecânica pelo tímpano e transformadas em sinais elétricos na cóclea. A figura à direita mostra um detalhe do ouvido médio (tímpano, martelo, bigorna e estribo) e a conexão com o ouvido interno (janela oval, cóclea).

Escutamos sons que ocorrem dentro de uma faixa limitada de intensidade, frequência e tempo. O limite da percepção de intensidade sonora começa a partir do chamado limiar da audição; o nível mínimo de variação de pressão acústica, onde o ouvido percebe a existência do som. A referência normalmente utilizada para o limiar de percepção é 20.10-6 N/m2, ou seja, Pref=20μPa (20.10-6Pa) de variação de pressão. Isto equivale a 0dB de pressão acústica (Sound Pressure Level, ou SPL). No limite superior tem-se o chamado limiar da dor, onde a intensidade sonora é tão grande que provoca sensação de desconforto ou dor. Acima desta intensidade, o som pode causar danos ao ouvido e consequentes perdas permanentes de audição. Este é normalmente definido como sendo em torno de 1 Pa, equivalente a 94dB SPL, dependendo da frequência e duração do sinal sonoro.<sup>5</sup> A escala de intensidade sonora é logarítmica.

 $<sup>^{\</sup>rm 5}$  Aparelhos áudio digital, como MP3 players e iPods podem facilmente ultrapassar 100dB SPL de emissão sonora. Quando utilizados

definida por L = 20.log10(Prms/Pref), onde Prms é o valor quadrático médio, ou RMS (Root Mean Square) da pressão sonora do sinal medido. A percepção da intensidade sonora é chamada de loudness. Apesar de relacionada com a intensidade sonora, o loudness é uma medida subjetiva e dependente de outros fatores tais como a sensibilidade do ouvinte, o tipo de sinal sonoro, sua envoltória ADSR (Ataque, Decaimento, Sustentação, e Repouso) e frequência. Apesar da cóclea ser a responsável pela captação do sinal sonoro, o sistema vestibular, responsável pelo equilíbrio, pode também ser estimulado por sons altos (acima de 90dB) especialmente aqueles com frequências graves, e pulsação rítmica. Por ser percebido pelo vestibular, tal estímulo é identificado deslocamento cinético corporal. O ouvinte que interpreta esta sensação como prazerosa, normalmente aprecia música com sons altos, com parciais de frequências graves e ritmados, que é facilmente agregado aos movimentos compassados do corpo, dança e deslocamentos rítmicos da cabeca. possivelmente afim de intensificar a sensação do referido estimulo sonoro (TODD et al., 2000).

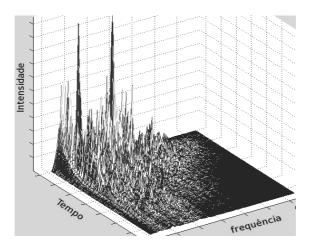
O limite de percepção da frequência sonora é relacionado ao formato em cunha do órgão de Corti, dentro da cóclea, que é sensível a frequências sonoras aproximadamente entre 20 Hz e 20 kHz (20.000 Hz). Para efeito de comparação, as frequências fundamentais das notas do piano, um dos instrumentos com maior extensão de escala musical, vão de 27,5 Hz para a primeira nota, o  $A_0$ , até 4.186 Hz, para a última nota, o  $C_8$ .6 A voz humana varia a frequência fundamental entre 80 Hz para baixos, até 1 kHz para sopranos. Os sinais acústicos, praticamente em sua totalidade, são sinais complexos, ou seja, formados por uma grande quantidade de parciais (ou formantes), que são os componentes básicos de qualquer sinal, incluindo o som. Parcial é uma onda sonora senoidal, cujos parâmetros variantes no tempo são amplitude, frequência e

frequentemente nessa faixa de intensidade sonora, podem causar danos permanentes à audição.

<sup>&</sup>lt;sup>6</sup> C<sub>4</sub> = dó central (na quarta oitava do teclado padrão do piano).

17

fase. O sinal sonoro mais simples possível é formado por um único parcial, cujo som equivale ao de um diapasão. Os parciais componentes de um som normalmente variam dinamicamente ao longo do tempo. Dependendo de seu comportamento, estes podem representar desde sons aperiódicos, como sons percussivos sem altura definida (p. ex., chocalho, surdo e recoreco), até sons praticamente periódicos, como os dos instrumentos musicais de altura determinada (p. ex., flauta, violão, piano). De um modo geral, a sensação de altura, ou tom, está relacionada à periodicidade do sinal sonoro. Sons aperiódicos não possuem uma altura definida, que possa descrever uma nota da escala musical. Estes são sons como os de aplausos, do quebrar das ondas do mar, explosões e disparos ou mesmo da fala humana (não cantada). Sons emitidos pela voz humana e diversos instrumentos de altura definida apresentam parciais de frequências múltiplas inteiras da fundamental f, ou seja, 2.f, 3.f, 4.f, etc. Esses parciais são chamados de harmônicos e sua ordenação em frequência (f, 2f, 3f, etc.) compõe a série harmônica. A claridade da altura está normalmente associada aos harmônicos no som, em ordem decrescente de intensidade, onde a fundamental é o harmônico de maior amplitude e sua frequência corresponde à frequência da altura do som. As componentes parciais podem ser representadas no domínio da frequência através Transformada de Fourier; nome dado em homenagem a Joseph Fourier, matemático que, no século XVII, provou que todo sinal (sonoro) pode ser descrito por uma somatória (finita, se o sinal for periódico) de senos e cossenos (parciais). A representação dos parciais no domínio da frequência é chamada de espectro, onde pode-se observar os parciais componentes de um som numa dada janela (intervalo) de tempo. A representação da variação do espectro ao longo do tempo é chamada de espectrograma. Este mostra a variação da intensidade dos parciais componentes de um som ao longo do tempo. A Figura 3 mostra o espectrograma do som do gongo, um instrumento de percussão.



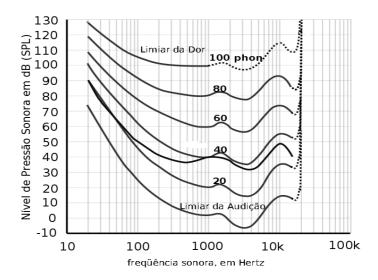
**Fig. 3:** Espectrograma do som de um gongo. A figura mostra a variação da intensidade dos parciais deste sinal ao longo da escala de tempo. Observa-se que no início do som existe uma imensa quantidade de parciais, que vão diminuindo de intensidade e desaparecendo ao longo do desenvolvimento deste som no tempo.

Para sinais sonoros simples (com apenas um parcial), a percepção da intensidade sonora varia desde abaixo de 0dB (para frequências em torno de 3 kHz), equivalente ao limiar da audição, até acima de 110dB, o limiar da dor, para frequências entre 100 Hz e 5 kHz. Estes limites foram inicialmente calculados pelos experimentos de Fletcher e Munson, onde foi demonstrado que, para sons simples (senoidais), a percepção da intensidade sonora é também dependente da sua frequência, neste caso, equivalente à sua altura (tom).

A unidade de *loudness* é chamada de *phon* e as curvas cujo *loudness* se mantém constante foram chamadas de curvas isofônicas; as curvas de igual intensidade perceptual sonora (*phon*), conforme mostradas na Figura 4, em sua versão atualizada, do padrão ISO 226:2003. Estes experimentos também ajudaram a mapear os limites de percepção sonora para sons simples, no que concerne a sua intensidade, em dB SPL e frequência em Hz, que determinam, como limites

18

aproximados da audição humana, intensidades sonoras variando entre 0 e 120 dB e frequências entre 20 e 20 kHz. É importante notar que os sons que ouvimos são praticamente sempre sons complexos, formados por uma imensa quantidade de parciais com intensidade, frequência e fase que variam independentemente ao longo do tempo, bem mais complexos que os sons senoidais utilizados no referido experimento.



**Fig. 4:** Curvas de igual intensidade perceptual sonora para sons simples.

Observa-se na figura que o ouvido é mais sensível à intensidade sonora dos parciais com frequências medianas, entre 1 kHz a 5 kHz, pois esta é a região de frequência relacionadas aos parciais gerados pela voz humana, o que garante um melhor entendimento das linguagens. Para sons complexos, a percepção da intensidade sonora é correspondente ao *loudness* resultante de todos os seus parciais. Porém, as ondas senoidais correspondentes aos

conhecido como mascaramento sonoro. Este ocorre tanto em intensidade como em frequência. Em termos de intensidade, um som mais intenso pode impedir (mascarar) que outros sons de menor intensidade, sejam ouvidos. Mascaramento em frequência ocorre quando dois parciais com frequências muita próximas acabam por não ser percebidos separadamente, devido aos limites de resolução da captação dos sinais sonoros na cóclea, onde sons muito próximos em frequência fazem vibrar uma mesma região do órgão de Corti e são desse modo percebidos como um único som.

parciais do som podem interagir entre si, criando o fenômeno

A audição também é sensível ao intervalo de tempo entre eventos sonoros, em pelo menos duas categorias: entre a geração de eventos simultâneos e entre a captação pelos dois ouvidos de um mesmo evento. Como dito antes, dois eventos subsequentes serão percebidos como eventos distintos somente se ocorrerem distanciados no tempo por um intervalo acima da persistência auditiva; cerca de 0,1s ou 10Hz. Comparativamente, um metrônomo padrão produz pulsos regulares entre 40 a 208 BPM (batimentos por minuto), o que equivale a frequências aproximadamente entre 0,7Hz e 3,5Hz. Num compasso quaternário (4/4) com quatro colcheias, este pulso de colcheias equivaleria à frequência de 14Hz. Acima de aproximadamente 20Hz passaríamos a não mais escutar as colcheias independentes, mas o todo, na forma de uma altura (tom) grave. Já o intervalo de tempo entre a captação de um mesmo evento sonoro pelos ouvidos é chamado de Diferença de Tempo Interaural, ou ITD, (Interaural Time Difference). Este é percebido como localização espacial horizontal, ou azimute da fonte sonora. Localização é calculada no primeiro estágio de processamento cerebral da informação sonora, no tronco encefálico, em duas regiões simétricas, cada uma conhecida como núcleo coclear, onde é conectado o nervo auditivo, vindo da cóclea de cada ouvido. Estudos demonstram que somos capazes de perceber a localização espacial de eventos sonoros com ITD acima de 0,00001s (10µs); considerado como sendo o limiar da localização sonora. No entanto, ainda não foi atribuída uma implicação cognitiva ou afetiva para este aspecto sonoro do sistema auditivo binaural, o que faz com que este seja aqui considerado como um aspecto perceptual do som.

20

De modo geral, podemos dizer que o sistema binaural de audição humana permite que sejamos sensíveis à presença e à variação da intensidade, periodicidade (frequência fundamental), composição espectral (presença e variação dos parciais formantes) e a localização sonora. Estas grandezas psicoacústicas constituem uma ampla gama de aspectos perceptuais, tais como aqueles que são descritos pela teoria musical (altura, intensidade, ritmo e timbre), os LLDs (descritos anteriormente, do padrão MPEG-7) e outros aspectos nãocontextuais da informação sonora, do domínio do tempo e da frequência. Através destes estabelecemos nossa primeira etapa de interação com o universo sonoro que compõe a música.

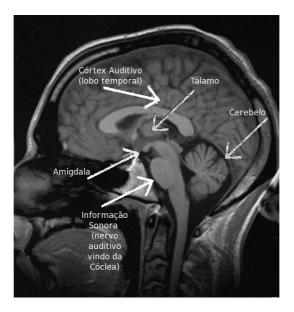
## A cognição musical

Quando estamos num ambiente tumultuado, como uma feira de domingo, percebemos uma imensa quantidade de fontes sonoras independentes, desconexas e em movimento. Escutamos ao nosso redor diversas vozes, simultaneamente falando ou conversando sobre assuntos distintos. Nossos ouvidos estão recebendo, traduzindo e enviando toda essa informação sonora para o cérebro pelo nervo auditivo, na forma de sinais elétricos. Apesar dessa informação perceptual estar emaranhada nos dois canais de recepção, que são os ouvidos, somos capazes, até certo ponto, de focalizar nossa atenção voluntariamente em uma única conversa, bem como mover nossa atenção de uma fonte sonora para outra, de acordo com nosso interesse, e desconsiderar o restante. Se, nesse ambiente sonoro tumultuado, alguém chama pelo nosso nome, especialmente se constatarmos que se trata de uma voz conhecida. imediatamente atenção nossa involuntariamente deslocada para esta pessoa. No entanto, nem sempre somos capazes de decidir pela informação sonora que iremos reter atenção. É o que ocorre quando escutamos uma geralmente de caráter repetitivo. necessariamente apreciamos (muitas vezes até repudiamos), e depois de inadvertidamente ouvi-la não conseguimos mais conter sua obstinada repetição em nossa mente. Este fenômeno cognitivo de repetição involuntária e compulsiva de um trecho musical foi denotado (em alemão) de Ohrwurm, cuja tradução é

22

"verme no ouvido" por quase sempre se referir a uma experiência bastante desprazerosa. Tenho observado pessoas idosas que se dizem incapazes de escutar um telefone tocando ao seu lado ou mesmo o som da televisão em alto volume. Porém, quando solicitadas a prestarem a atenção, são capazes de ouvir um som simples e sutil de um diapasão.

Fenômenos como estes evidenciam a existência de um tratamento mental ativo e sofisticado da informação sonora pela percepção auditiva, que privilegia certas informações enquanto desconsidera outras, de acordo com critérios dinâmicos, não necessariamente voluntários. A informação sonora vinda da cóclea pelo nervo auditivo, na forma de impulsos elétricos, chega inicialmente numa região do tronco encefálico, ou cerebral, chamada de núcleo coclear. Esta é a primeira região cerebral conectada ao nervo auditivo. O núcleo coclear é localizado na ponte do tronco cerebral que se conecta ao cerebelo onde é processada a informação de localização espacial (azimute, ou localização horizontal) juntamente com a sensação de equilíbrio, também captada no ouvido interno, pelo sistema vestibular. Em seguida, a informação é enviada ao tálamo, uma estrutura cerebral constituinte do sistema límbico, que se conecta ao hipocampo (que regula a memória de longa-duração), e à amígdala (que regula a agressividade e o medo no indivíduo). Do tálamo, a informação processada segue ao córtex auditivo, a primeira região do córtex cerebral, localizado no lobo temporal, que recebe informação sonora. O córtex é composto pela região superficial, mais externa do cérebro, constituído pelas últimas camadas desenvolvidas durante a evolução biológica, e está presente apenas nos cérebros de mamíferos. O córtex regula sofisticadas funções cognitivas, como atenção, memória, linguagem e consciência. A Figura 5 mostra as regiões do cérebro humano descritas acima relacionadas e processamento da informação sonora e sua interpretação como informação musical.



**Fig. 5:** As regiões do cérebro humano relacionadas ao processamento da informação sonora.

Existem diversas anormalidades da cognição musical detectadas em indivíduos portadores de danos em algumas regiões cerebrais específicas. Por exemplo, a amusia, ou "surdez tonal" é a inabilidade de reconhecer variações de pitch, relacionado à capacidade conhecida como "ouvido relativo", ou seja, a faculdade de reconhecer e discriminar o intervalo de altura musical entre sons tonais. Em sua forma congênita, esta parece estar associada ao desenvolvimento anormal da região cerebral do córtex auditivo (PERETZ, 2008). Outra desabilidade causada pela lesão das duas regiões do lobo temporal, onde está localizado o córtex auditivo, é conhecida como agnósia musical. Nela, o indivíduo é capaz de reconhecer a fala, a linguagem, sons da natureza, até mesmo a letra de canções que o paciente conhecia antes da lesão, mas não é capaz de reconhecer suas melodias, ou seja, o indivíduo perde a capacidade de percepção tonal (altura), mas mantém a capacidade de percepção rítmica (tempo) (PERETZ, 1996). Indivíduos com lesões nos dois hemisférios da amígdala

apresentam inabilidade em reconhecer músicas de contexto assustador ou triste, ou seja, estes são incapazes de reconhecer o significado musical relacionado à sensação de medo. No entanto, tais indivíduos são capazes de reconhecer o contexto de alegria ou felicidade (nas músicas com tal contexto). Isto sugere que a amígdala é responsável pelo reconhecimento de emoções relacionadas ao medo (GOSSELIN, 2007). Arritmia auditiva é o nome dado à incapacidade de manter ou reconhecer um padrão rítmico ou pulsação rítmica estável. Estudos sugerem que tal capacidade está relacionada a lesões no lobo temporal direito, mais especificamente, no córtex auditivo direito (WILSON, 2002).

Uma habilidade bastante conhecida entre os músicos é chamada de "ouvido absoluto". Indivíduos com tal capacidade conseguem reconhecer e lembrar de tons (altura das notas da escala musical cromática) sem a necessidade de uma referência externa (como um instrumento musical ou diapasão). Apesar da controversa impossibilidade de se adquirir voluntariamente tal habilidade, estudos sugerem que esta capacidade é de caráter puramente cognitivo e não perceptual. Indivíduos com ouvido absoluto que foram examinados não apresentaram diferenças significativas na estrutura auditiva ou cerebral relacionada à audição, nem tampouco foram encontradas diferenças de sensibilidade ou acuracidade da percepção de outros aspectos musicais. Existe, no entanto, uma maior incidência de indivíduos com ouvido absoluto que tiveram formação musical na infância, bem como entre falantes de línguas tonais, como o mandarim e cantonês. Também existe uma maior incidência dessa habilidade na população de indivíduos cegos de nascença, bem como de portadores de autismo.

Na música, o processo cognitivo interpreta os aspectos perceptuais sonoros e os interpreta como informação musical. Do repertório de interpretações, algumas são instintivas enquanto outras são aprendidas. Chama-se aqui de instinto, a capacidade inata (não-aprendida) de um indivíduo apresentar espontaneamente um comportamento, normalmente relacionado às reações básicas de manutenção de seu conforto ou sua sobrevivência, como a agressividade, o apetite, ou o interesse sexual. Já as capacidades aprendidas são aquelas de

24

caráter cultural, que o indivíduo necessita desenvolver através da observação ou comunicação com outros indivíduos, para cultivá-las, antes de se tornarem proficientes em sua performance e posteriormente ensiná-las a outros indivíduos. Instinto e aprendizado parecem cumprir funções similares; a de garantir a sobrevivência do indivíduo (normalmente por instinto) e de sua população (normalmente por aprendizado). Alguns estudos sugerem que um determinado padrão comportamental que seja constantemente aprendido pelos indivíduos de uma população, com o passar do tempo (ao longo de suas gerações) tende a se tornar um padrão comportamental instintivo. Este é conhecido como efeito Baldwin (HURON, 2006). Por essa teoria, processos cognitivos (de aprendizado) cumprem também a função ontogênica (evolução individual) de tornar a adaptação do indivíduo mais rápida às mudanças súbitas da pressão ambiental imposta sobre sua população, de modo a proporcionar a adaptação filogênica (evolução social). Podemos dizer que a maioria dos aspectos perceptuais da música, conforme descritos anteriormente, são instintivos, uma vez que não precisam ser aprendidos pelo indivíduo (como é o caso da percepção da presença e variação da intensidade sonora, altura, localização espacial horizontal, e timbre). Em termos de aspectos cognitivos, aqueles relacionados ao processamento límbico parecem ser instintivos enquanto os relacionados ao processamento no córtex são aprendidos. Como será visto adiante, estes correspondem aos caminhos rápido (instintivo) e lento (aprendido) de processamento simultâneo da informação sonora, responsáveis ao afeto musical.

O processo cognitivo musical se beneficia dessas duas categorias de processamento; uma instintiva e a outra aprendida. Diversos estudos foram conduzidos sobre a distinção cognitiva instintiva entre a escuta de música tonal (com um centro tonal perceptível) e atonal (sem um centro tonal perceptível). Em um deles, foram analisados neonatos com 2 dias após seu nascimento, de pais surdos (o que supostamente implicaria em neonatos praticamente sem exposição à música, durante toda sua gestação) e pais com audição normal, onde os bebês haviam sido normalmente expostos a estímulos musicais durante sua gestação. Observou-

se que todos os neonatos demonstraram preferência pela consonância à dissonância (NOBUO, 2006). Pode-se comparar a predileção ao tonalismo com a preferência de uma criança por doces. Instintivamente, esta prefere o sabor que identifica o alimento de maior carga energética. Porém, com seu aprendizado e crescimento, este desenvolverá também o gosto por outros sabores mais elaborados. Até mesmo os que antes eram intragáveis, podem com o passar do tempo suplantar em sua ordem pessoal de preferência.

Aspectos cognitivos da música tratam do material musical cujo significado depende do seu contexto. Este é mediado pela memória e consequente identificação de similaridades e contrastes com O material musical anteriormente escutado, seja este microestruturais (aspectos identificados durante a escuta de uma mesma peça musical) ou macroestruturais (aspectos identificados na peça que a relacionam a um gênero musical). Esta pode ser estudada com base na psicologia da forma, ou Gestalt, que considera o todo cognitivo como maior (mais significativo) que a soma de seus aspectos, ou seja, o significado não pode ser linearmente reduzido e representado por seus aspectos componentes. A Gestalt trata com a identificação de objetos, ou seja, ao reconhecimento imediato que temos de objetos visuais ou seu correspondente auditivo, os eventos sonoros. Por exemplo, quando encontramos numa multidão um rosto conhecido, o seu reconhecimento parece ocorrer imediatamente, sem que exista um processo de busca num banco de dados mental de feições e expressões. Reconhecemos o rosto mesmo que este rosto esteja diferente de quando o conhecemos (mais envelhecido, penteado diferente, etc.). Existem quatro princípios básicos da identificação de objetos. No caso de um evento musical, temos: 1) Estruturação; capacidade de entender, em determinadas circunstâncias, como uma estrutura única, um conjunto de eventos distintos (este é o caso da identificação de uma melodia, linha de baixo, harmonia, ritmo, etc.). 2) Segregação; processo cognitivo que identifica um evento perceptualmente mais saliente, em contraste a outros eventos menos destacados (também chamados de eventos de fundo). Um exemplo disto observa-se quando, numa obra sinfônica, facilmente identificamos a melodia de um flautim, mesmo que toda a

26

orquestra esteja também tocando simultaneamente. 3) Pregnância; a tendência de identificarmos primeiramente (ou mais claramente) estruturas que sejam mais simples e regulares. É o caso de padrões rítmicos simples que são mais fáceis de se identificar que padrões complexos, poli-rítmicos ou ad-libitum. 4) Constância; identificação de continuidade nas variações de aspectos perceptuais entre eventos consecutivos de modo a garantir a compreensão de que estes se referem a um mesmo contexto. Exemplo disso é quando escutamos um veículo com alto-falantes passar por nós irradiando uma música. Escutamos grandes variações de aspectos perceptuais, tais como intensidade, altura (resultante do efeito *Doppler*) e localização espacial. No entanto, ainda assim conseguimos facilmente identificar que se trata do mesmo evento sonoro.

A predição de aspectos cognitivos pode atualmente ser realizada através de descritores cognitivos. Estes são modelos computacionais descrevem aspectos que contextuais. simulando a cognição musical humana. Exemplos de tais descritores são relacionadas à identificação de existência (clareza), complexidade, variação (relativa e absoluta) e similaridade de aspectos musicais contextuais, tais como pulso rítmico, padrão rítmico, tonalidade, modo, melodia, linha de baixo, demais linhas melódicas (p. ex., contrapontos), harmonia, etc. Diversos descritores dessa categoria foram simulados e apresentaram resultados interessantes, obtendo correlação suficiente com medidas similares humanas, feitas com grupos de ouvintes (FORNARI et al., 2009).

## A emoção evocada pela música

É comum escutarmos uma música e sermos surpreendidos pela evocação de memórias episódicas, contendo fatos, situações, locais e lembranças de pessoas. Normalmente sentimos emoções associadas a essas memórias, muitas vezes involuntárias e até mesmo inesperadas. Sabe-se que a música é bastante eficiente em despertar emoções nos ouvintes. Por isso tem sido utilizada por outras formas artísticas a fim de criar ou intensificar uma prosódia afetiva. É o caso da canção, a união da poesia com a música, que vêm sendo parceiras desde tempos imemoriais, produzindo obras

28

religiosas, profanas, folclóricas, populares e eruditas. Essa associação parece ser bastante eficiente especialmente porque une a capacidade de significação semântica da linguagem com a significação afetiva da música. Desde o seu início, o cinema também tem se associado à música, dos primeiros filmes mudos, que eram "acompanhados" por música ao vivo, nos teatros de exibição, até a sua forma atual, onde são elaboradas complexas trilhas sonoras, muitas vezes chegando a utilizar orquestras sinfônicas para sua gravação. O cinema também utiliza da sonoplastia, arte responsável pela criação dos ambientes sonoros relacionados ao discurso filmográfico. Observa-se que, normalmente, a trilha sonora trata de intensificar a prosódia afetiva enquanto a sonoplastia intensifica a prosódia perceptual e cognitiva. Similar à música, a poesia e o cinema também são formas de expressão artística que se desenvolvem ao longo do tempo. Todas apresentam uma prosódia, se bem que a música parece ser mais atrelada ao andamento temporal. Pode-se ler um texto em prosa, ou declamar uma poesia, em diferentes andamentos ou métricas sem alterar substancialmente seu contexto afetivo. no entanto. pequenas alterações do andamento musical podem gerar mudanças significativas da performance musical. Aliás, são muitas vezes essas pequenas variações que distinguem uma interpretação musical boa de uma interpretação excelente. Esta característica do afeto musical é um dos fatores que a torna tão eficiente no seu papel de evocar emoções. Tal como no drama e na comédia, das artes cênicas, na música a precisão da prosódia do interprete é fundamental ao efetivo significado do contexto afetivo.

Estudos nas áreas da psicologia e neurociência tem sido realizados no sentido de se entender como a música evoca emoções (SLOBODA et al., 2001). A literatura menciona três modelos principais de emoção musical: 1) o modelo categórico, originado dos trabalhos de Ekman (1992) e Juslin et al. (2003) que descreve a música em termos de uma listagem de categorias de emoções distintas; 2) o modelo dimensional, originado nos trabalhos de Russell (1980) e Laukka et al. (2005) onde é proposto que as emoções podem ser decompostas em emoções básicas e ortogonais (que podem ser tratadas como variáveis independentes), num sistema de

coordenadas; e 3) modelo do processo componente, de Scherer et al. (2001), que descreve a constatação da emoção musical de acordo com a situação de sua ocorrência e o presente estado emocional do ouvinte. Por seu aspecto reducionista, diversos estudos têm utilizado um modelo dimensional com duas dimensões, ou seja, um modelo bidimensional de emoção para mapear as emoções descritas pelo modelo categórico. As duas dimensões utilizadas normalmente são valência e atividade. Valência representa a emoção de satisfação, e se estende do insatisfeito ao satisfeito. Atividade representa a emoção de excitação e se estende do sonolento ao excitado. Tais emoções são consideradas ortogonais. Através destas coordenadas emocionais, é possível descrever uma série de emoções experimentadas pela mente humana. A Figura 6 mostra o modelo bi-dimensional de afeto, também conhecido por modelo circumplexo de Russel. Observa-se nesta figura que categorias de emoção estão dispostas aproximadamente circular. Estas são dadas como pontos no plano (espaço euclidiano R2). Exemplificando, a emoção alegre é correspondente à atividade e valência positivas. Já a emoção tenso é correspondente à atividade positiva e valência negativa, enquanto que a emoção triste é formada por valores de atividade e valência negativas.

Do mesmo modo que algoritmos foram desenvolvidos para a predição de aspectos perceptuais (modelados por LLDs), modelos computacionais, para o cálculo de aspectos afetivos (relacionados às emoções evocada pela música), vem também sendo desenvolvidos por grupos acadêmicos concentrados em torno da área de conhecimento atualmente referida por MIR (*Music Information Retrieval*). Entre outras pesquisas, podemos citar o desenvolvimento de um modelo computacional para classificação de gênero musical (TZANETAKIS *et al.*, 2002); o estudo de aspectos gestuais da performance musical (LEMAN *et al.*, 2003); e o desenvolvimento de descritores acústicos de altonível (contextuais) para a extração de aspectos afetivos musicais (WU *et al.*, 2006; GOMEZ *et al.*, 2004).



Fig. 6: O modelo circumplexo de Russel28.

Em seu estudo sobre o desenvolvimento dinâmico de emoção musical, (SCHUBERT, 1999) utilizou o modelo circumplexo para medir, em relação ao tempo, a variação de emoção musical constatada por um grupo de ouvintes sobre diversas pecas do repertório clássico tradicional. As dimensões emocionais utilizadas foram as mesmas mostradas na Figura 6 (atividade e valência). Foram apresentados dois modelos lineares para cada peca musical; um para descrever atividade e outro para valência. Posteriormente, Korhonene et al. (2006) utilizou as mesmas medidas comportamentais para criar dois modelos gerais; um para a dimensão atividade e outro para a dimensão valência. Ambos os estudos não obtiveram bons resultados na previsão de valência, provavelmente porque ambos utilizaram apenas descritores de baixo-nível (perceptuais) em seus modelos. Finalmente, em estudo realizado por Fornari (2008) utilizando o mesmo corpo de dados dos estudos anteriores, foi desenvolvido um modelo com descritores de alto-nível (densidade de eventos, complexidade harmônica, brilho, claridade de pulso, repetição, articulação,

claridade de tonalidade e modo). Como este, foi possível prever adequadamente a valência para a mesma peça musical. No entanto, apesar de elucidar a possibilidade de predição das dimensões ligadas à constatação da emoção com descritores de alto-nível, estes não tratam necessariamente da emoção evocada pela música, mas apenas da constatação do ouvinte em relação à intenção emotiva de uma peça musical do repertório tradicional.

Sabe-se que alguns aspectos da emoção evocada estão associados às variações de sinais biológicos involuntários, também chamados de biossinais, tais como a variação da resistência cutânea (Galvanic Skin Response - GSR), a variação de batimento cardíaco através de eletrocardiograma (ECG) ou fotopletismografia (PPG) e a variação do ritmo respiratório. Em termos de afeto musical, tais biossinais podem ser utilizados para descrever tanto a variação de estados emocionais evocados de curta duração (affect), como os de longa duração (mood) (THAYER, 1989). Affects podem estar relacionados a curtos trechos musicais, ou trechos intramusicais (da mesma música) dentro do intervalo de tempo do "agora musical". Moods são geralmente criados por períodos mais longos de escuta musical, relacionados ao tamanho total de uma obra ou performance musical, como a duração média de uma sinfonia, um show musical ou a duração de um CD (de 60 a 90 minutos de duração). Seus efeitos emotivos persistem por longos períodos e podem ser verificados na análise das variações de ritmos biológicos do indivíduo, tal como o ciclo circadiano (MOORE-EDE, 1982).

Na estética da música erudita ocidental do século XX é comum nos depararmos com a noção de que a informação sonora que compõe a música, não possui, em si própria, elementos que diretamente correspondam às emoções evocadas nos ouvintes pela música, mas que esta representa um encadeamento de contextos advindos da organização de sons, que o ouvinte eventualmente associa com emoções relacionadas a fatos ocorridos em sua história de vida (memória). Este conceito vem dos estudos de Eduard Hanslick, que fundamentou a teoria estética musical opondo-se à teoria dos afetos (HANSLICK, 1957). No século XIX, Christian Friedrich Michaelis desenvolveu a teoria dos afetos como uma

32

extensão à teoria estética de Kant sobre música. Kant afirmava que a música em sua forma pura (sem texto) é uma arte nãorepresentativa, e como tal pode exprimir o belo livre, ou seja, aquele que não é associado a qualquer outra representação. Em linhas gerais, a teoria dos afetos afirma que a música é uma arte que tem como propósito exprimir sentimentos através da modulação dos sons. A teoria estética de Hanslick desvinculou a arte musical deste compromisso exclusivo com a expressão do belo, ou da transmissão de emoções prazerosas aos ouvintes. Disso veio a noção estética vigente no século XX que a música deve primar pela criação de estruturas musicais cuja engenhosidade e complexidade (similar à arquitetura) seria eventualmente cógnita (percebida, compreendida e apreciada) ao ouvinte. Esta teoria concentra todo o significado estético musical na obra em si, desconsiderando sua cognição, ou seja, a descodificação desta informação musical pelo ouvinte. Sem cognição não há comunicação e a obra passa a ser incógnita. O significado da música está necessariamente atrelado à sua compreensão e esta à comunicação efetiva dos conceitos estruturados na composição, mediados pela performance e adequadamente decifrados pela percepção, cognicão afetividade do ouvinte.

A afetividade, ou seja, a evocação de emoções, pode ser estudada também sob uma perspectiva evolutiva. Darwin e Ekman exemplificaram isto através da análise de expressões faciais de emoções comuns, como alegria, surpresa, medo, raiva, etc. Tais expressões são monomórficas, ou seja, são comuns a todos os indivíduos da espécie humana, não importando sua origem etno-sócio-cultural (GRIFFITHS, 1997) o que leva a supor que essas expressões cumprem um papel atávico na comunicação humana do estado emocional. (MEYER, 1957) explica que emoções são evocadas quando a tendência de resposta é inibida, ou seja, no momento em que um comportamento é interrompido, seja por um impedimento externo ou deliberadamente restrito pelo indivíduo (p. ex., um motorista preso num engarrafamento, ou um indivíduo tentando conter um acesso de riso). Nestas situações são geradas as emoções mais intensas. Assim, a lei do afeto de Meyer é sumarizada como "emoções são evocadas quando a tendência a uma ação é inibida". Já o modelo afetivo de JamesLange (de William James e Carl Lange) afirma que emoções resultam da experiência de mudanças corporais. Exemplificando, sentimos medo porque experimentamos taquicardia, e não o contrário. Este controverso modelo foi relegado por muito tempo mas atualmente vem contando com evidências neurocientíficas a seu favor. Modelos funcionalistas, como o de Nico Frijda, afirmam que emoções evoluíram para um propósito ou função específica, como o de garantir o bemestar ou sobrevivência do indivíduo.

Um modelo de emoção musical evocada de acordo com tais premissas evolutivas foi proposto por (HURON, 2006). Este é chamado de ITPRA (Imagination, Tension, Prediction, Reaction Segundo este modelo, (algum)as emoções and Appraisal). evocadas pela música advêm de uma estratégia evolutiva de gerar sensações de conforto ou satisfação, quando o indivíduo consegue estabelecer previsões acertadas a eventos externos (p. ex., emoções satisfatórias, como amor, alegria, admiração), bem como o de punir previsões incorretas com sensações de desconforto (emoções desconfortáveis, como medo, raiva, tristeza). Este sistema de previsão de eventos, incluindo os eventos musicais, é processado simultaneamente por dois caminhos neurológicos distintos. Um lento, processado pelo córtex auditivo, relacionado ao conhecimento aprendido (reconhecimento, lembrança, associação, similaridade). O outro caminho é de ação rápida; processado pelo sistema límbico, e relacionado às reações instintivas, de sobrevivência, do tipo fight, flight or freeze (lutar, fugir ou paralisar). A Figura 7 demonstra este modelo, na linha do tempo de processamento mental para a ocorrência de um evento musical. Inicialmente tem-se a imaginação, que reúne o conjunto de expectativas com relação ao evento que está por ocorrer. Este pode ser composto de macroestruturas, como o gênero musical, o estilo da performance, a composição, bem como as estruturas microestruturas, como a melodia, o encadeamento harmônico, o andamento rítmico, etc. A imaginação estabelece as expectativas do ouvinte em relação aos eventos musicais que poderão ocorrer (futuros). A tensão é a reação anterior a ocorrência de um evento que já é praticamente certo que irá ocorrer, que permite que o ouvinte se prepare (p. ex., a tensão imediatamente anterior a uma nova repetição de um refrão

musical). Enquanto a tensão aumenta, a imaginação diminui, proporcionalmente à diminuição da probabilidade ocorrência de outros eventos inesperados, durante a tensão. No instante da ocorrência do evento, a tensão desaparece subitamente. Tem-se imediatamente 0 processamento simultâneo de duas ações: a predição, relacionada processamento lento, e a reação, ao processamento rápido acima). A reação é relacionada (ambos descritos processamento límbico e envolve ações automáticas, não processadas no córtex, que fomentam ações básicas e imediatas, relacionadas a garantir a sobrevivência do indivíduo. Do mesmo modo, a predição é iniciada no córtex auditivo. Este correlaciona o evento ocorrido com os modelos mentais esperados na Imaginação e estabelece a similaridade entre estes, prevendo sua implicação. Na medida em que predição e reação diminuem, aumenta a constatação, que avalia o grau de acerto da predição; se esta foi correta, em relação à imaginação e tensão, bem como se a reação foi adequada ao tipo de evento musical.



Fig. 7: O modelo de emoção musical evocada ITRPA, de Huron.

O modelo ITRPA tenta responder através da psicologia das expectativas uma questão fundamental: Por que a música é uma obsessão humana? É fato que a vasta maioria das pessoas escuta música porque as faz se sentirem bem. Com o advento de equipamentos portáteis de áudio digital e seu custo cada vez menor, pessoas estão escutando música mais do que nunca. O acesso à música, que há pouco menos de um século atrás era restrito a locais e ocasiões especiais, sem qualquer possibilidade de retenção dessa informação sonora, que não fosse por sua lembrança, agora pode ser gravada, reproduzida,

copiada e escutada em praticamente qualquer ambiente e circunstância. Por sua capacidade de representar emoções monomórficas (como as expressões faciais emotivas), pelo menos em grandes grupos sociais, a música tem sido organizada em repertórios particulares e muitas vezes utilizada pelos ouvintes como forma de induzir ou estimular um estado afetivo específico (p. ex., uma música ritmada para realizar um exercício físico, uma música suave para relaxar ou induzir o sono, etc.). Um fato interessante é que as músicas que evocam negativas (como tristeza 011 solidão) necessariamente evocam nos ouvintes sensações desconforto. Na verdade, estas são muitas vezes escutadas por indivíduos que compartilham de um estado emocional similar ao expresso pela música ouvido. Muitas pessoas vivendo em ambientes urbanos e ruidosos com excesso de poluição sonora demonstram preferência por músicas populares acusticamente similares (p. ex., punk, rap, funk), da mesma maneira que muitas pessoas vivendo no campo preferem músicas que se encaixem no contexto deste ambiente sonoro (p. ex., sertanejo, moda-deviola, caipira). Esta é uma constatação puramente empírica e ainda carece de rigoroso estudo para se sustentar como teoria, no entanto, parece que a música pode também vir a ser utilizada como uma forma de adequar o indivíduo à sua realidade ou a uma mudança de realidade, seja esta do seu estado emocional ou ambiente, similar a uma das funções do sonho; a de consolidação da memória semântica (PAYNE, 2004). Um indivíduo que tenha passado por uma desilusão amorosa pode desenvolver uma predileção maior por canções românticas, que descrevam situações similares a sua, ao invés de preferir músicas de caráter alegre, como marchas ou polcas. Um motorista, que diariamente enfrenta o transito caótico da metrópole, pode preferir escutar *heavy-metal* à música barroca. O aspecto afetivo da música pode cumprir a função de ajudar na regulação do estado emocional do indivíduo, propiciando uma melhor inserção em seu ambiente sócio-cultural.

BALDWIN James Mark. A new factor in evolution. *American Naturalist*, v. 30 (1896), p. 441-451, 536-553.

THAYER, Robert E. *The biopsychology of mood and arousal*. New York: Oxford University Press, 1989.

EKMAN, Paul. An argument for basic emotions. *Cognition & Emotion*, v. 6, n. 3/4 (1992), p. 169–200.

FORNARI, José. Predição de aspectos da emoção constatada em música por descritores cognitivos musicais. *Anais do SIMPEMUS 5*. Curitiba: DeArtes – UFPR, 2008.

FORNARI, José; EEROLA, Tuomas. The pursuit of happiness in music: retrieving valence with contextual music descriptors. In: YSTAD, Sølvi; KRONLAND-MARTINET, Richard; JENSEN, Kristoffer (orgs.). Computer music modeling and retrieval: genesis of meaning in sound and music. Proceedings of the 5th International Symposium CMMR 2008, p. 119-133. Berlim / Heidelberg: Springer, 2009. Série "Lecture Notes in Computer Science", v. 5493.

FORNARI, José; MANZOLLI, Jônatas; SHELLARD, Mariana. O mapeamento sinestésico do gesto artístico em objeto sonoro. *Opus*, v. 15, n. 1 (jun. 2009), p. 69-84.

GOMEZ, Emília; HERRERA, Perfecto. Estimating the tonality of polyphonic audio files: cognitive versus machine learning modeling strategies. *Proceedings of the 5th ISMIR*. Barcelona, October 2004.

GOSSELIN, Nathalie; PERETZ, Isabelle; JOHNSEN, Erica; ADOLPHS, Ralph. Amygdala damage impairs emotion recognition from music. *Neuropsychologia*, v. 45, n. 2 (2007), p. 236-244.

GRIFFITHS, Paul E. What emotions really are: the problem of psychological categories. Chicago: Chicago University Press, 1997.

HANSLICK, Eduard. *The beautiful in music*. Trad. Gustav Cohen. Indianapolis: Bobbs-Merrill, 1957.

HURON, David. Sweet anticipation: music and the psychology of expectation. Cambridge, MA: MIT Press, 2007.

JACOBS Lawrence; FELDMAN, Martin; BENDER, Morris. The persistence of visual or auditory percepts as symptoms of irritative lesions of the cerebrum of man. *Journal of Neurology*, v. 203, n. 3 (set. 1972), p. 211-218.

JAMES, William. The principles of psychology. New York: Cosimo, 2007.

JUSLIN, Patrik N.; LAUKKA, Petri. Communication of emotions in vocal expression and music performance: different channels, same code? *Psychological Bulletin*, v. 129, n. 5, (set. 2003), p. 770-814.

KORHONEN, Mark; CLAUSI, David; JERNIGAN, M. Ed. Modeling emotional content of music using system identification. *IEEE Transactions on Systems, Man and Cybernetics*. v. 36, n. 3 (jun. 2006), p. 588-599.

LAUKKA, Petri; JUSLIN, Patrik; BRESIN, Roberto. A dimensional approach to vocal expression of emotion. *Cognition & Emotion*, v. 19, n. 5 (ago. 2005), p. 633-653.

LEMAN, Marc; VERMEULEN, Valery; DE VOOGDT, Liesbeth; TAELMAN, Johan; MOELANTS, Dirk; LESAFFRE, Micheline. Correlation of gestural musical audio cues and perceived expressive qualities. In: *Gesture-Based Communication in Human-Computer Interaction. 5th International Gesture Workshop*, p. 40-54. Berlim / Heidelberg: Springer, 2004. Série "Lecture notes in artificial intelligence", v. 2915.

MANJUNATH, B. S.; SALEMBIER, Philippe; SIKORA, Thomas (orgs.) *Introduction to MPEG-7: multimedia content description interface*. Nova York: Wiley & Sons, 2002.

MEYER, Leonard B. *Emotion and meaning in music*. Chicago: Chicago University Press, 1956.

MOORE-EDE, Martin C.; SULSZMAN, Frank M.; FULLER, Charles A. *The clocks that time us: physiology of the circadian timing system*. Cambridge, MA: Harvard University Press, 1982.

NARA Takahiro; GOTO, Noburu; HAMANO Shin-Ichiro; OKADA, Akira. Morphometric development of the human fetal auditory system: inferior collicular nucleus. *Brain and Development*, v. 18, n. 1 (1996), p. 35-39.

NOBUO Masataka. Preference for consonance over dissonance by hearing newborns of deaf parents and of hearing parents. *Developmental Science*, v. 9, n. 1 (2006), p. 46-50.

PAYNE, Jessica D.; NADEL, Lynn. Sleep, dreams, and memory consolidation: the role of the stress hormone cortisol. *Learning & Memory*, v. 11 (2004), p. 671–678.

PERETZ, Isabelle. Can we lose memory for music? A case of music agnosia in a nonmusician. *Journal of Cognitive Neuroscience*, v. 8, n. 6 (1996), p. 481-496.

PERETZ, Isabelle. *Musical disorders: current directions in psychological science*, v. 17, n. 5 (2008), p. 329-333.

RUSSELL, James A. A circumplex model of affect. *Journal of Personality and Social Psychology*, v. 39 (1980), p. 1161-1178.

SCHERER, Klaus R.; ZENTNER, Marcel R. Emotional effects of music: production rules. In: SLOBODA, John; JUSLIN, Patrik (orgs.), *Music and emotion: theory and research*, p. 361-392. Oxford: Oxford University Press, 2001.

SCHUBERT, Emery. Measuring emotion continuously: validity and reliability of the two-dimensional emotion space. *Australian Journal of Psychology*, v. 51, n. 3 (1999), p. 154-165.

SLOBODA, John; JUSLIN, Patrik (orgs.): *Music and emotion: theory and research*. Oxford: Oxford University Press, 2001.

TODD, Neil; CODY, Frederick. Vestibular responses to loud dance music: a physiological basis for the 'rock and roll threshold'? *Journal of the Acoustical Society of America*, v. 107, n. 1 (2000), p. 496–500.

TZANETAKIS, George; COOK, Perry. Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, v. 10, n. 5 (2002), p. 293-302.

WILSON, Sarah J.; PRESSING, Jeffrey L.; WALES, Roger J. Modelling rhythmic function in a musician post-stroke. *Neuropsychologia*, v. 40, n. 88 (2002), p. 1494-1505.

WU, Tien-Li; JENG, Shyh-Kang. Automatic emotion classification of musical segments. *Proceedings of the 9th ICMPC* - International Conference on Music Perception & Cognition, Bologna, 2006.

# Técnicas fundamentais para instrumentos computacionais de síntese e processamento de áudio

VICTOR LAZZARINI

este capitulo, exploraremos os ciemes construção de instrumentos de síntese e processamento construção de instrumentos de síntese e processamento que o este capítulo, exploraremos os elementos básicos para a de sinais para aplicações musicais. Para isso, assumimos que o leitor tenha alguma familiaridade com as linguagens de programação C ou C++. Também será importante o domínio de alguns dos rudimentos da teoria de sinais digitais, como o conceito de amostras e do teorema da amostragem, e de princípios matemáticos, como fundamentos de trigonometria e álgebra. Os elementos discutidos neste texto formam a base de todos os sistemas de processamento de sinais musicais: osciladores, envelopes, filtros e retardos. Com estes é possível o desenvolvimento de instrumentos musicais (WANDERLEY, 2010) que implementem várias das técnicas estabelecidas de síntese e processamento. Estas incluem, por exemplo, as sínteses aditiva, subtrativa e por modulação de frequência (DODGE e JERSE, 1985); o processamento granular (BARREIRO e KELLER, 2010); os modelos físicos (guias de ondas) (SMITH, 2004) e de formantes (LAZZARINI e TIMONEY, 2009); e os efeitos de reverberação, chorus, flanger, transposição, e outros processos similares (ROADS, 1996).

#### **Osciladores**

O primeiro componente essencial para a síntese de áudio é o oscilador (DODGE e JERSE, 1985; MOORE, 1990), usado primariamente para a geração de sinais periódicos, como por exemplo uma onda senoidal. Osciladores podem ser usados em várias aplicações e são um elemento genérico de sistemas de processamento de áudio.

Partindo de primeiros princípios, vamos discutir algumas questões básicas relacionadas a osciladores. Começando com ondas senoidais, observamos que estas podem ser sintetizadas com um simples uso da função sin():

a[n] = sin(n\*2\*pi\*f/sr);

Isso gerará uma senóide com frequência f usando a taxa de amostragem¹ sr (onde pi representa a constante  $\pi$  em precisão dupla, double pi). Este exemplo funcionará para alguns casos simples, mas não em todas as aplicações. O problema principal é que ele falhará nos casos em que a frequência for variável, por exemplo se quisermos gerar um glissando.

A função sin() espera como parâmetro, um ângulo, ou melhor, uma fase.² Quando passamos a ela uma expressão que depende de um valor de frequência, teremos que transformá-la primeiramente em uma fase, o que de fato é feito no exemplo acima. A fase muda continuamente com o tempo (que é representado pelo número da amostra n),  $2\pi \times f \times n/sr$  (n\*2\*pi\*f/sr). O problema é que se a frequência varia de amostra para amostra, o valor da fase não vai ser calculado devidamente. Para resolver isso, temos que calcular a fase levando em conta que a frequência é equivalente à diferença entre dois valores consecutivos da fase (ou melhor, ela é a primeira derivada da fase³), por isso integramos a frequência para obter a fase correta:

<sup>1</sup> Ver glossário.

ver grossurio

<sup>&</sup>lt;sup>2</sup> Fase, ou ângulo, é o argumento de funções trigonométricas como o seno e cosseno. Se variarmos regularmente a fase destas no tempo produziremos uma onda senoidal.

 $<sup>^3</sup>$  A frequência, que determina o número de ciclos da onda senoidal por unidade de tempo (por exemplo, segundos), tem um valor instantâneo que é definido pela diferença de dois ângulos (fases) relacionados com amostras consecutivas (digamos n e n-1). Se somarmos estes valores instantâneos, obteremos a fase corrente (ou atual) da onda.

```
fase = 2*pi*f/sr.
a[n] = sin(fase);
fase += 2*pi*f/sr;
```

#### Tabelas de onda

Em geral, para construirmos um oscilador genérico, não vamos querer nos limitar ao uso direto de funções como sin(), pois é muito mais interessante se pudermos usar funções periódicas arbitrárias. Por isso, usaremos a técnica de tabelas de ondas (Fig.1). Estas são blocos de memória que contém um ciclo de uma forma de onda arbitrária para ser usada em síntese. Em termos de C++, serão vetores de números reais (float, double) de qualquer tamanho. Na verdade, o tamanho da tabela vai determinar a precisão do processo, e com isso a qualidade do sinal gerado. Uma tabela de senóide pode ser gerada assim:

```
for(int i=0; i < tamanho; i++)
     tabela[i] = sin(i*2*pi/tamanho);</pre>
```

A saída do oscilador é determinada primariamente pelo processo de consulta à tabela, que é basicamente, a leitura de um valor do vetor que contém a onda, e em seguida a sua multiplicação pela amplitude:

```
s[n] = a*tabela[(int)indice];
```

Pelo fato de que o índice usado para ler a tabela ser float, temos que transformá-lo em um valor inteiro. Este processo é chamado de truncagem. O índice será constantemente atualizado para que um sinal seja produzido.

Como vimos acima, a frequência fundamental do oscilador depende do incremento da fase. No caso de um oscilador de tabelas, este é geralmente chamado de *incremento de amostragem*, e é calculado de acordo com a frequência, o tamanho da tabela e a taxa de amostragem:

```
incr = f*tamanho/sr;
indice += incr;
```

Finalmente, é preciso que o índice seja mantido dentro da escala correta (0 até N-1, onde N é o tamanho da tabela). Para isso aplicamos a operação de módulo generalizado<sup>4</sup> (para o oscilador também poder funcionar com frequências e incrementos negativos):

```
while(indice >= tamanho) indice -= tamanho;
while(indice < 0) indice += tamanho;</pre>
```

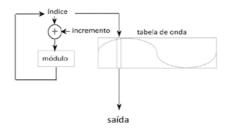


Fig. 1: O algoritmo do oscilador

#### Processando blocos de áudio

Para se implementar um processador de áudio como o oscilador, é mais eficiente trabalhar-se com blocos de amostras, ao invés de processá-las individualmente. Neste caso, se criarmos uma função para codificar o oscilador, esta será chamada apenas uma vez para cada bloco de amostras, e isso será computacionalmente mais eficiente. Como consequência, estaremos criando duas taxas de processamento: uma, de amostra a amostra, equivalente a taxa de amostragem de áudio; e a outra, de bloco a bloco, mais lenta, uma taxa de amostragem de sinais de controle. Voltaremos ao assunto um pouco mais a frente.

<sup>&</sup>lt;sup>4</sup> O módulo aqui é definido como  $a \equiv b \pmod{n}$  se a - b é um múltiplo inteiro de n. O módulo também é definido para números negativos, por exemplo  $-11 \equiv 1 \pmod{10}$ 

```
/** oscilador por truncagem.
  saida: bloco de saida \n
  amp: amplitude \n
  freq:
          frequencia \n
  tab: tabela de funcao \n
  ind: indice \n
  tam:
         tamanho da table \n
  bloco: tamanho do bloco \n
  sr: taxa de amostragem\n
                        \n
  return: 1
* /
float osc(float *saida, float amp, float freq,
              float *tab, float *ind,
              int tam=def_len, int bloco=def_vsize,
              long sr=def_sr);
```

As constantes def\_len, def\_vsize e def\_sr, relativas ao tamanho da tabela, tamanho de bloco e taxa de amostragem, foram definidas previamente (1024, 256 e 44100). Esta é a implementação do oscilador:

Note-se que produzimos um bloco completo no vetor saída, e também produzimos como resposta a primeira amostra da tabela, o que facilitará o uso da função como controle. Importante também é notar que temos que manter o valor do índice armazenado externamente à função, pois este terá que

manter o seu valor entre chamadas a ela. Por isso usamos ponteiros (\*ind). Para se usar este oscilador, temos que colocálo em um laço de processamento, por exemplo:

```
for(int i=0; i < dur; i++){
  osc(buffer,amp,freq,onda,&ndc);
  soundout(psf,buffer);
}</pre>
```

Sinais de controle

Introduzimos, acima, o conceito de taxa e sinais de controle. Por razões de eficiência, podemos decidir usar alguns processos nesta taxa mais lenta, produzindo sinais que não serão usados diretamente como áudio, mas apenas para controlar ou modular parâmetros de síntese (como por exemplo a frequência, produzindo um vibrato). O exemplo abaixo demonstra este conceito. Note-se que um oscilador produz o sinal de áudio e outro produz um sinal de controle, que contém apenas uma amostra para cada bloco de amostras do sinal de áudio. O oscilador de controle que modula a frequência possue bloco = 1 e taxa de amostragem = def\_cr (equivalente a def\_sr/def\_vsize, lembrando que estas constantes são definidas previamente):

```
for(int i=0; i < dur; i++){
  osc(buffer,amp,freq +
      osc(&cs,10.f,5.f,onda,&ndc2,def_len,1,def_cr),
      onda,&ndc);
  soundout(psf,buffer);
}</pre>
```

Gerando tabelas de onda

Provavelmente o melhor método para se gerar tabelas de onda para osciladores é usando a série de Fourier:

```
a = amps ? amps[i] : 1.f;
w = (i+1)*(n*2*pi/tam);
tab[n] += (float) (a*cos(w+fase));
}
normalizar_tab(tab, tam);
return tab;
}
```

No código acima, usamos uma função externa para normalizar a tabela,<sup>5</sup> pois a adição de componentes usando a série de Fourier pode fazer com que a amplitude passe do máximo (1).<sup>6</sup> Note-se também que adicionamos um ponto extra ao final da tabela, para facilitar o processo de interpolação (veja abaixo). Esta função geral pode ser usada para construir ondas 'clássicas' como a dente-de-serra, quadrada, triangular, etc. Por exemplo:

```
float* serra_tab(int harms, int tam){
float *amps = new float[harms];
for(int i=0; i < harms; i++) amps[i] = 1.f/(i+1);
float *tab = fourier_table(harms,amps,tam, -0.25);
delete[] amps;
return tab;
float* quadra_tab(int harms, int tam){
float *amps = new float[harms];
memset(amps, 0, sizeof(float)*harms);
for(int i=0; i < harms; i+=2)amps[i] = 1.f/(i+1);
float *tab = fourier_table(harms,amps,tam, -0.25);
delete[] amps;
return tab;
float* triang_tab(int harms, int tam){
float *amps = new float[harms];
memset(amps, 0, sizeof(float)*harms);
for(int i=0; i < harms; i+=2)amps[i]= 1.f/((i+1)*(i+1));
float *tab = fourier_table(harms,amps,tam);
delete[] amps;
return tab;
```

 $<sup>^{5}</sup>$  Neste caso, normalizar significa colocar os valores da tabela entre -1 e 1.

<sup>&</sup>lt;sup>6</sup> É mais conveniente que a tabela tenha valor máximo = 1, pois então poderemos controlar a amplitude do sinal mais precisamente (usando o parâmetro do oscilador).

## Osciladores por interpolação

Interpolação é uma operação usada em muitas aplicações em áudio digital, consistindo em preencher valores não existentes previamente, quando necessário. Por exemplo, se decidirmos obter uma amostra de uma tabela, fracionalmente entre duas posições, teremos que interpolar para obter o valor (próximo do) correto. Esta é a base do oscilador por interpolação, que é uma evolução sobre o oscilador por truncagem.

O método mais simples de interpolação é o linear. Se um indíce p cair entre pontos  $x_1$  e  $x_2$ , cujos valores na tabela são  $y_1$  e  $y_2$ , então, linearmente, temos

$$y = y_1 + (y_2 - y_1) \times (p - x_1)$$
 (1)

Aqui  $x_1$  e  $x_2$  são índices inteiros consecutivos e p é o índice fracional<sup>7</sup>. Interpolação linear não é mais que o uso de uma função de primeiro grau:

$$y = cx + d \tag{2}$$

onde  $d = y_1$ ,  $c = y_1 - y_2$  e  $x = p - x_1$ .

O código-fonte abaixo implementa um oscilador por interpolação linear e inclui também um controle de fase (0-1). Note-se que esperamos que a tabela contenha um ponto a mais para que o processo de interpolação funcione:

<sup>&</sup>lt;sup>7</sup> Ou seja, o índice real entre dois valores integrais. Por exemplo, para p=3.5, temos  $x_1=3$  e  $x_2=4$ .

```
// interp. linear
frac = pos - (int)pos;
a = tab[(int)pos];
b = tab[(int)pos+1];
output[i] = amp*(a + frac*(b - a));
    *indice += incr;
    while(*indice >= length) *indice -= tam;
    while(*indice < 0) *indice += tam;
}
return *output;
}</pre>
```

# **Envelopes**

Envelopes podem ser gerados por vários métodos. O princípio básico é se ter uma certa curva que moldará a amplitude (ou outro parâmetro) do som. Estas podem ser geradas por uma função que interpole entre dois pontos, produzindo diretamente um sinal de controle. Também podemos gerar tabelas contendo estas curvas para serem lidas por osciladores.

A mais simples das tabelas de envelope pode ser gerada interpolando-se entre dois pontos, usando os mesmos princípios discutidos acima. Para se gerar uma tabela com N pontos, linearmente, entre 0 e 1, podemos usar:

```
for(int i =0; i < N; i++)
    tab[i] = (float)i/N;</pre>
```

Geralmente, podemos dizer:

```
tab[i] = inicio + i*(fim - inicio)/N;
```

que também pode ser implementada assim:

```
tmp = inicio;
incr = (fim - inicio)/N;
for(int i =0; i < N; i++) {
  tab[i] = tmp;
  tmp += incr;
}
```

Este código gerará um segmento linear. Tabelas com multiplos segmentos podem ser construídas com base neste fragmento. Envelopes lineares funcionam bem para controles simples. No entanto para sermos mais precisos em termos perceptuais, especialmente com frequências, precisamos usar interpolação exponencial:<sup>8</sup>

$$y(x) = y_1 \times (y_2 / y_1)^x$$
 (3)

Mas é preciso cuidado, y<sub>1</sub> e y<sub>2</sub> devem ser positivos, e não podem ser zero.<sup>9</sup> Para se implementar este processo, temos

```
for(int i =0; i< N; i++)
    tab[i] = start*pow(fim/inicio, (double) i/N);</pre>
```

#### Ou também:

```
tmp = inicio;
mult = pow(fim/inicio, 1./N);
for(int i=0; i < N; i++){
  tab[i] = tmp;
  tmp *= mult;
}</pre>
```

# Exemplos

Aqui oferecemos dois exemplos de funções geradoras de tabelas de envelope com múltiplos segmentos:

```
float* linear_tab(int brkpts, float* pts, int tam){
  float inicio,fim,incr,*tab = new float[tam+1];
  for(int n=2; n < brkpts*2; n+=2){
    inicio = pts[n-1];
    fim = pts[n+1];
    incr = (fim - inicio)*1.f/(pts[n]-pts[n-2]);</pre>
```

<sup>&</sup>lt;sup>8</sup> A nossa percepção de frequências é, em geral, logarítmica; percebemos razões entre frequências e não diferenças. Por isso para se produzir glissandos ou outras mudanças que soem mais naturais, precisamos usar envelopes exponenciais.

<sup>&</sup>lt;sup>9</sup> Para evitar divisão por zero e porque a expressão só é definida para valores positivos.

A função abaixo mostra um exemplo que gera tabelas com envelopes exponenciais:

Podemos usar estas geradoras de tabelas em programas com osciladores. Por exemplo, criamos uma tabela de 400 pontos com três segmentos (envelope trapezoidal):

```
float pts[6] = {0,0,100.f, 1.f, 400.f, 0.f};
env = linear_tab(3, pts, 400);
```

Agora então podemos utilizá-la com um oscilador para controlar a amplitude de um som:

```
oscc(buffer,
    osc(&out,amp,1/dur,env,&ndc2,400,1,def_cr),
    freq,wave,&ndc);
```

O oscilador lerá a tabela de envelope somente uma vez durante a duração do som (note que sua frequência é 1/dur, onde dur é a duração do som em segundos).

# Geradores de envelope

Uma desvantagem de se usar tabelas de envelope é que a duração dos segmentos será sempre ligada à duração total do som. A solução alternativa é empregar geradores de envelope, que produzirão um sinal com uma certa duração fixa. Pelo fato de que estes sempre funcionarão como controladores, vamos implementá-los gerando sinais de controle. Similarmente aos osciladores, estes processadores precisarão manter um indíce, externamente ao código da função. No entanto, o uso deste aqui é como um contador de posição de tempo (e por isso deve ser um número inteiro).

Um gerador de linha é o exemplo mais simples, produzindo um sinal que varia de um ponto a outro em um dado intervalo de tempo. O que acontece depois é indeterminado, mas podemos fazer com o que o sinal se limite ao último valor alcançado (pos 2):

```
float line(float pos1, float dur, float pos2, int *cnt,
float cr){
    int durs = (int) (dur*cr);
        if((*cnt)++ < durs)
            return pos1 + *cnt*(pos2-pos1)/durs;
        else return pos2;
}</pre>
```

Usando o mesmo princípio, podemos ter um gerador de envelope exponencial:

O gerador de envelope mais comum em muitos sintetizadores é o ADSR (fig.2). Possuindo quatro estágios, ataque, decaimento, sustentação e repouso (ou relaxamento), é geralmente implementado com segmentos Tipicamente, em sintetizadores, o ADSR é controlado por um sinal externo que o 'liga' e 'desliga'. Quando ligado o envelope percorre o seu curso até o estágio da sustentação, e somente entra em seu repouso quando é desligado. No entanto, aqui como estamos determinando a duração do som, vamos colocar os quatro segmentos dentro da duração total. Isso quer dizer que o período de sustentação será calculado pela diferença dur (at + dt + rt). O começo do último estágio do envelope, repouso, é calculado pela diferença entre a duração total e o tempo de repouso (rt).

Neste algoritmo, durações são em segundos e por isso precisam ser primeiramente convertidas em número de amostras. O segmento de ataque vem primeiro:

```
if(*cnt <= at) a = *cnt * (maxamp/at);</pre>
```

onde a amplitude cresce de 0 a maxamp em at segundos. Em seguida, temos o decaimento, onde a amplitude passa da maxamp a sus em dt segundos:

O contador ent é incrementado a toda amostra, e então observamos se há tempo sobrando para o período de sustentação:

```
else if(*cnt <= (dur - rt)) a = sus;
```

Se não for o caso, vamos direto para o último estágio, repouso, que trará a amplitude para 0 ao fim:

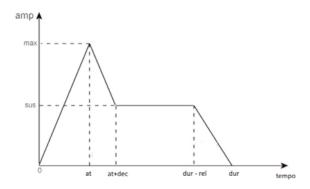


Fig: 2: O gráfico do envelope ADSR

# Aqui está então o ADSR, completo:

```
float adsr(float maxamp, float dur, float at, float dt,
            float sus, float rt, int *cnt, float cr){
float a;
// converter o tempo
at = at*cr;
dt = dt*cr;
rt = rt*cr;
dur = dur*cr;
if(*cnt < dur) { // se tempo < dur
    // ataque
    if(*cnt <= at) a = *cnt * (maxamp/at);</pre>
    // decay
    else if(*cnt <= (at+dt))</pre>
     a = ((sus - maxamp)/dt)*(*cnt - at) + maxamp;
     // sus
    else if(*cnt <= (dur - rt))
      a = sus;
     // release
    else if(*cnt > (dur - rt))
      a = -(sus/rt)*(*cnt - (dur - rt)) + sus;
else a = 0.f;
// incrementar o contador de tempo.
(*cnt)++;
return a;
```

## Um exemplo completo

Antes de tentarmos colocar os elementos acima em um programa completo, devemos decidir onde colocar o sinal digital produzido. O que faremos é apresentar uma interface geral de entrada e saída de áudio, que pode ser implementada de várias formas. Aqui a usaremos para produzir arquivos de áudio, mas é possível implementá-la novamente para, por exemplo, utilizar uma entrada e saída de som em tempo real.

Criaremos três tipos de funções: 1) para abrir arquivos (que pode ser implementada para abrir uma saída para placa de som); 2) para escrever ou ler as amostras (e estas também podem ser implementadas para usar uma placa de som); e 3) para fechar arquivos (ou entrada/saída de som). Abaixo mostramos as declarações destas funções:

```
typedef SSOM void*;
SSOM soundout_open(char* nome, int chans=def_chans,
              float sr=def_sr);
SSOM soundin_open(char* nome, int chans=def_chans,
              float sr=def_sr);
void soundout_close(SSOM psf_out);
void soundin_close(SSOM psf_out);
/** saída de som
   psf_out: descritor da saida
   buffer: buffer (interlacado se multicanal)
   vecsize: tamanho do buffer em frames
   return: numero de frames colocados no destino
int soundout(SSOM psf_out, float *buffer,
              int vecsize=def_vsize);
/** entrada de som
   psf_in: descritor da entrada
   buffer: buffer (interlacado se multicanal)
   vecsize: tamanho do buffer em frames
   return: numero de frames recebidos da fonte
int soundin(SSOM psf_out, float *buffer,
              int vecsize=def vsize);
```

 $<sup>^{\</sup>rm 10}$  Todos os sinais produzidos aqui serão digitais, definidos como discretos em tempo e em amplitude.

Usando por exemplo, a biblioteca *libsndfile*<sup>11</sup>, pode-se facilmente implementar estas funções. Mas estas devem ser tratadas como 'caixas-preta', pois não importa como são implementadas, mas o que fazem (levar o som ao seu destino).

O nosso exemplo completo abaixo mostra o uso de osciladores e envelopes:

```
#include <stdio.h>
#include <stdlib.h>
#include <snd_defs.h>
/** sintetizador simples com envelopes
Gera um som com uma onda dente-de-serra controlada
      por envelopes \n
       \n
       sint arquivo.wav amp freq(Hz) dur(secs)
* /
int main(int argc, char** argv) {
  SSOM *psf;
  float *buffer;
  int smps, cnt1=0, cnt2=0;
  float dur, amp, freq, *onda, ndc=0;
  if(argc == 5){
   amp = (float) atof(argv[2]);
   freq = (float) atof(argv[3]);
   dur = (float)atof(argv[4]);
   smps = (int) (dur*def_cr);
   // alocar memoria
   buffer = new float[def_vsize];
   onda = serra_tab();
   // abrir arquivo ou saida de som
   if(!(psf = soundout_open(argv[1]))){
    printf("erro ao abrir arquivo ou saida\n");
   exit(-1);
   for(int i=0; i < smps; i++){
    osci(buffer,
       amp*adsr(1.f,dur,0.05f, 0.1f, 0.7f, 0.2f, &cnt1),
       expon(freq,dur/2,freq*2, &cnt2),
       onda, &ndc);
      soundout(psf,buffer);
    }
```

Libsndfile é uma biblioteca especializada escrita por Erik de Castro Lopo para entrada e saída de áudio para arquivos de vários formatos, http://www.mega-nerd.com/libsndfile

#### **Filtros**

Filtros são processadores de sinal usados para transformações timbrísticas (STEIGLITZ, 1995). Sua função principal é esculpir o espectro do som, atenuando ou amplificando componentes em certas bandas de frequência.

Implementamos filtros combinando sinais e suas cópias com retardo, de várias maneiras. Existem duas famílias básicas de filtros, em relação à forma em que combinamos os sinais<sup>12</sup>:

- (1) Filtros que usam (somente) sinais de entrada atrasados, chamados de Resposta Finita a Impulso (RFI). Estes filtros são sempre estáveis.
- (2) Filtros que usam sinais de saída atrasados chamados de Resposta Infinita a Impulso (RII) (que também podem incluir sinais de entrada atrasados). Estes filtros podem ser instáveis.

Em alguns filtros digitais, observaremos sinais com retardos, ou *retardos*, de apenas uma amostra. A *ordem* do filtro é determinada pelo retardo máximo usado na operação. Por

<sup>&</sup>lt;sup>12</sup> Filtros RFI não são tão usados em síntese como os RII. Uma das razões para isso é devida ao fato de que filtros RII podem ter suas características (parâmetros) modificadas dinamicamente, o que é mais difícil em filtros RFI. Também o tamanho e número de operações necessárias para se produzir um efeito expressivo é menor em filtros RII.

exemplo, se o filtro usa retardos de duas amostras, este será de segunda ordem<sup>13</sup>.

Filtros são definidos por suas equações, que mostram os retardos usados no filtros e os coeficientes (ganhos) associados à eles, por exemplo:

$$y(n) = x(n) + a_1 x(n-2) - b_1 y(n-1) - b_2 y(n-2)$$
 (4)

onde n é um índice denotando tempo em amostras; y(n) é o sinal de saída e x(n), o sinal de entrada;  $a_1$  é o coeficiente associado com um retardo de duas amostras do sinal de entrada;  $b_1$  e  $b_2$  são os coeficientes associados, respectivamente, com os retardos de uma e duas amostras do sinal de saída.

A resposta de frequência de um filtro determina como este altera um sinal de entrada, em termos de amplitude e fase em várias frequências. A resposta de amplitude determina como cada frequência é amplificada ou atenuada. Quanto à esta, geralmente classificamos os filtros em quatro tipos (Fig. 3):

- passa-banda: passa ou amplifica componentes em uma certa banda, atenuando aquelas fora desta.
- rejeita-banda: o contrário do tipo anterior, rejeita ou atenua frequências em uma banda;
- passa-baixa: passa ou amplifica componentes abaixo de uma certa frequência, chamada frequência de corte.
- passa-alta: passa ou amplifica componentes acima da frequência de corte.

argumento (para mais detalhes veja STEIGLITZ, 1995).

 $<sup>^{13}</sup>$  A ordem do filtro é também relacionada com a ordem do polinômio que determina a sua função de transferência. No caso do filtro definido pela eq.4, esta é definida por  $_{H(z)=rac{1-a_{z}z^{-2}}{1+b_{z}z^{-1}+b_{z}z^{-2}}}$ . Aqui as variáveis  $z^{-1}$  e

 $z^{-2}$  correspondem aos retardos de uma e duas amostras. A resposta de frequência é esta função com  $z=e^{j\omega}$ , quando calculada em termos de várias frequências f com  $\omega=2\pi ft$ . A resposta de amplitude é a magnitude desta função complexa, e a reposta de fase, o seu

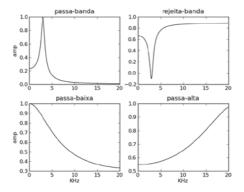


Fig. 3: Os quatro tipos comuns de filtro

Finalmente, a *resposta de fase* determina os retardos (ou mudanças de fase) que o filtro aplica a cada frequência do sinal de entrada.

#### Ressonadores

Ressonadores são um tipo básico de filtros RII que formam um componente essencial de sistemas de síntese de som. Examinaremos nesta seção os princípios de programação destes tipos de filtro. A equação básica do resonador é:

$$y(n) = ax(n) - b_1 y(n-1) - b_2 y(n-2)$$
 (5)

Como se vê, este filtro é de segunda ordem, com coeficientes  $a_0$ ,  $b_1$  e  $b_2$ . A resposta de frequência do filtro será determinada pelos seus coeficientes. Por isso, podemos determinar os coeficientes de acordo com uma resposta de frequências que desejamos.

O ressonador é um filtro cuja resposta de amplitude é do tipo passa-banda com dois parâmetros característicos: (i) frequência central; (ii) banda. Os coeficientes  $b_1$  e  $b_2$  serão determinados de acordo com estas equações (onde f é a

frequência central, B, a banda e sr, a frequência de amostragem, em Hz):

$$R = 1 - \pi (B/sr) \tag{6}$$

$$b_1 = -[4R^2/(1+R^2)]cos(2\pi f/sr)$$
(7)  
$$b_2 = R^2$$
(8)

$$b_2 = R^2 \tag{8}$$

A banda do filtro pode também ser determinada usando o fator de qualidade Q = f/B.

O volume de saída do ressonador tende a variar com a frequência central e a banda, e também com o sinal de entrada. Com ressonâncias estreitas, onde a banda é muito pequena, é possível que o filtro produza um sinal muito alto que poderá causar distorção. Para evitar este problema, podemos usar o coeficiente  $a_0$  que atenuará o sinal. Para se fazer com que o sinal na frequência central tenha amplitude 1, podemos usar:

$$a_0 = (1-R^2)\sin(2\pi f/sr)$$
 (9)

Um dos problemas que vemos em ressonadores é associado a uma deformação em sua resposta de amplitudes quando a frequência central está perto de 0 Hz ou de sr/2 (metade da frequência de amostragem, ou frequência de Nyquist). A atenuação abaixo (ou acima) da frequência central não é tão eficiente nestes casos. Uma solução proposta é adicionar-se um sinal de entrada com um retardo de duas amostras (SMITH e ANGELL, 1982; STEIGLITZ, 1994), cujo efeito é criar um ponto de atenuação em 0Hz e na Nyquist.

$$y(n) = a_0 x(n) - a_2 x(n-2) - b_1 y(n-1) - b_2 y(n-2)$$
 (10)

Uma função em C++ implementando o ressonador tem esta forma. Note-se que usamos o mesmo espaço de memória para a entrada e a saída (o sinal processado substitui o original):

```
float resonador(float* sig, float freg, float b,
              float *del,int vecsize,float sr){
double r, rsq, rr, costh, a;
```

A segunda versão deste filtro usando o retardo de duas amostras do sinal de entrada é implementada usando-se a equação 10, com:

$$a_o = l - R \tag{11}$$

$$a_2 = R(I - R) \tag{12}$$

Algumas vezes, quando usamos filtros, é importante que controlemos o volume do sinal cuidadosamente. Este é o caso, por exemplo, quando conectamos dois ou mais filtros em série (ou seja, a saída de um alimentando a entrada de outro). Para isso, podemos construir um processador que balanceie o volume do sinal de saída de acordo com um sinal de entrada. A

ideia é estimar a amplitude (RMS)<sup>14</sup> dos dois sinais, e modificar o sinal de saída de acordo com a razão entrada/saída.

Para se descobrir a amplitude RMS de um sinal, um dos métodos mais comuns envolve a retificação (obtendo os valores absolutos das amostras) e a filtragem usando-se um filtro simples de passa-baixa, com uma frequência de corte por volta de 10Hz. Esta determina, por proporção inversa, o intervalo de tempo em que a média do sinal é calculada (para se obter a amplitude). Demonstramos essas ideias no código-fonte abaixo. Dois sinais são usados, um a ser modificado e o outro usado como comparação. Obtemos o valor absoluto e o passamos para o filtro que usa a seguinte equação (de primeira ordem):

$$y(n) = x(n) (1 + C) - y(n-1)C$$
 (13)

onde C é o coeficiente do filtro, que para ter a forma de passa baixa é definido como:

$$\sqrt{(2-c-2\sigma\tau f/ss)^2 r \phi 1} - 2 + c - 2\sigma\tau f/ss) \tag{14}$$

Note-se que a retificação é combinada com o filtro em uma só linha de código. Cada sinal tem o seu filtro separado, e

<sup>&</sup>lt;sup>14</sup> RMS (ou *root-mean-square*) é um método de se obter a amplitude média de um sinal, levando em conta a energia de um sinal. É definido pelo processo de se obter a raiz quadrada da média do quadrado do sinal (neste caso das amostras).

no final usamos os sinais de saída do filtro para obtermos a razão entre a entrada (sinal comparativo) e a saída. Este operador de balanço é então usado pós-filtragem, com um comparativo pré-filtragem.

## Um programa de exemplo

Aqui demonstramos as ideias introduzidas acima em um programa completo:

```
#include <stdio.h>
#include <stdlib.h>
#include <snd defs.h>
/** sintese subtrativa com envelopes
       Gera um som com uma onda dente de serra filtrada
       por um ressonador com envelopes de amplitude e de
       frequencia central.
       filtro sndfile.wav amp freq(Hz) dur(s)
* /
int main(int argc, char** argv) {
 SSOM *psf;
 int smps, bytes = sizeof(float)*def_vsize,cnt1=0,
      cnt2=0;
 float dur, amp, freq, *buffer, *wave, *comp, ndx=0;
 if(argc == 5){
   amp = (float) atof(argv[2]);
  freq = (float) atof(argv[3]);
  dur = (float)atof(argv[4]);
  smps = (int) (dur*def_cr);
   // alocar memoria
  buffer = new float[def vsize];
  comp = new float[def_vsize];
   float del[2]=\{0.f, 0.f\}, dell[2]=\{0.f, 0.f\};
  wave = saw_table(30);
   // abrir arquivo ou saida
   if(!(psf = soundout_open(argv[1]))){
   printf("erro ao abrir o arquivo ou saida\n");
   exit(-1);
   for(int i=0; i < smps; i++){
      osci(buffer,
       adsr(amp,dur,0.05f, 0.1f, amp*0.7f, 0.2f, &cnt1),
       freq, wave, &ndx);
     memcpy(comp, buffer, bytes);
      resonador(buffer, expon(freq*8,dur,freq*4,
                              &cnt2, 50, del);
      balanc(buffer, comp, del1);
```

#### Processadores de retardo

Muitos dos efeitos de áudio digital são baseados em um tipo de processador comum: o retardo, que, em termos gerais, produz um atraso de tempo em um sinal de entrada. O tempo de retardo nestes processos é normalmente muito maior que o que vimos acima em filtros, mas também pode variar de poucos milisegundos a alguns segundos.

Retardos de pequenas dimensões, como no caso visto em filtros, são implementados por uma simples estrutura de fila primeiro-dentro primeiro-fora<sup>15</sup> (PDPF ou FIFO em inglês), com amostras sendo copiadas de uma posição de memória a outra, cada vez que consumimos e produzimos uma amostra. No entanto, em retardos mais longos, este procedimento não é recomendado. Neste caso usamos um *buffer*<sup>16</sup> circular, que mantém as amostras no mesmo local em que foram escritas e move as posições de leitura e escrita para simular a passagem das amostras pelo retardo (fig.4). Ao chegar ao fim do bloco de memória usado pelo buffer, estas posições ou ponteiros são trazidos de volta ao começo, procedendo então circularmente.

 $<sup>^{15}\,</sup>$  Neste caso, a primeira amostra a entrar na fila é a também a primeira a sair.

 $<sup>^{16}\,</sup>$  Um bloco de memória de computador; em termos de programação, um vetor.

Por isso, o uso de *buffers* circulares é muito eficiente e elegante: só necessitamos manter uma conta das posições de leitura e de escrita, e o número de operações é independente do tamanho do retardo.

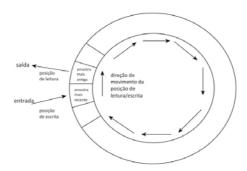


Fig. 4: Uma representação esquemática do buffer circular

#### Retardos fixos

Para várias aplicações, nós apenas necessitamos de retardos com uma duração fixa. Neste caso, o *buffer* circular tem uma forma bem simples. Aqui mostramos um exemplo deste processador, que impõe um retardo de um certo tempo a um sinal. A função em C++ requer que um bloco de memória externa seja alocado para o *buffer* circular, e que o índice que contém a posição de leitura/escrita também exista externamente (como nos exemplos anteriores de osciladores). O algoritmo é baseado em três estágios:

- Ler o buffer do retardo na posição atual para produzir a amostra de saída
- 2. Escrever a amostra de entrada naquela posição
- Avançar a posição por uma unidade e checar se não chegamos ao final do buffer (em cujo caso devemos retornar ao começo)

```
float s;
  dt = (int) (tpoddel*sr);
  for(int i=0; i < vecsize; i++){
    s = del[*p];
    del[*p] = sig[i];
    sig[i] = s;
    *p = (*p != dt-1 ? *p+1 : 0);
  }
  return *sig;
}</pre>
```

O tempo de retardo (tpoddel) determina o tamanho do buffer (del) que será usado, e deve ser menor ou igual ao espaço de memória alocado para o retardo. Este processador pode ser usado para um eco simples, se combinado com o sinal original. A aplicação típica de retardos fixos é para a construção de reverberadores e de câmeras de eco. Para estes usos, vamos estudar dois tipos de unidades básicas de processamento, o filtro pente e o filtro passa-todas. Reverberadores podem ser construídos conectando-se várias unidades destas em diversas combinações<sup>17</sup>.

## Filtros pente

O filtro pente é basicamente o retardo simples discutido acima com uma linha de re-alimentação conectando a saída do retardo à sua entrada. O nome 'pente' vem de sua resposta de amplitude, que tem a forma de um pente invertido (ou seja com cerdas para cima), com picos de amplificação espaçados uniformente a 1/t Hz, onde t é o tempo de retardo.

A quantidade de realimentação é determinada pelo parâmetro g (o ganho de realimentação ou feedback):

$$g = 0.001^{\frac{\tau}{RVT}} \tag{13}$$

<sup>&</sup>lt;sup>17</sup> Para maiores detalhes sobre reverberadores, consulte (DODGE e IERSE, 1985) e (BOULANGER, 2000).

onde RVT é o tempo total de reverberação (o tempo que levará um sinal de entrada a cair até -60dB), e t é o tempo de retardo, como já dito acima. O valor de g deve ser menor que um e positivo. Modificando-se o código da função retardo(), temos:

Filtros pente são usados em uma variedade de aplicações. Sozinhos podem ser empregados para a geração de ecos múltiplos, ou então como câmera de ressonância (com retardo curto e ganho de realimentação perto de 1.0). Em reverberadores, vários deles são conectados em paralelo com as suas saídas alimentando filtros passa-todas em série.

# Filtros passa-todas 18

Estes componentes de reverberadores são construídos similarmente aos filtros pente, a diferença sendo que eles possuem uma linha ligando a entrada do filtro diretamente á sua saída, onde se aplica o mesmo valor do ganho de realimentação, mas com sinal invertido. Este é o código:

<sup>&</sup>lt;sup>18</sup> O termo passa-todas (allpass) é relacionado com o fato de que este filtro, em seu estado estável, passa todas as frequências sem alterar relativamente as suas amplitudes, ao contrário dos filtros-pente, que têm uma resposta a amplitudes bem característica. No entanto, este filtro passa-todas altera as fases do sinais de entrada, já que sua resposta de fase não é linear.

```
float s;
  dt = (int) (tpoddel *sr);
  for(int i=0; i < vecsize; i++){
    s = retardo[*p];
    retardo[*p] = sig[i] + s*gain;
    sig[i] = s - g*sig[i];
    *p = (*p != dt-1 ? *p+1 : 0);
}
  return *sig;
}</pre>
```

Vários filtros passa-todas são geralmente conectados em série para aumentar a densidade de ecos em efeitos de reverberação.

#### Retardos variáveis

Se pudermos variar o tempo de retardo dinamicamente, uma série de efeitos será possível: *flanger, chorus,* vibrato, *doppler,* tranposição de alturas, etc. Para se implementar retardos variáveis, quatro fatores terão que ser considerados:

- A posição de escrita avança uma unidade a cada nova amostra. A posição de leitura então é calculada em relação à de escrita, menos o retardo desejado.
- A posição de escrita poderá ultrapassar o limite do bloco de memória (buffer), e neste caso teremos que usar uma operação de módulo para trazê-la de volta ao valor correto.
- 3. A posição de leitura poderá cair entre duas posições de memória. Neste caso teremos que interpolar as amostras lidas da linha de retardo para obter um sinal sem muita distorção (como no caso dos osciladores).
- 4. Quando usarmos interpolação linear, precisamos de duas amostras adjacentes no buffer circular. Um caso especial, que teremos que observar em nosso algoritmo, ocorre quando as duas amostras não são contíguas em memória (fim e começo do buffer circular).

Como um exemplo do uso de retardos variáveis, implementaremos aqui um efeito de *flanger*, que é baseado no filtro-pente, mas com a possibilidade de se variar dinamicamente o tempo de retardo. Neste caso, usaremos interpolação linear:

```
float flanger(float *sig, float tpoddelv, float fdb,
              float maxdel, float *retardo, int *p,
               int vecsize, float sr){
 int mdt,rpi;
 float s, rp, vdt, frac, next;
 vdt = tpoddelv*sr;
 mdt = (int) (maxdel*sr);
 if(vdt > mdt) vdt = (float) mdt;
 for(int i=0; i < vecsize; i++){
   rp = *p - vdt;
   rp = (rp >= 0 ? (rp < mdt ? rp : rp - mdt) :
              rp + mdt);
   rpi = (int) rp;
   frac = rp - rpi;
   next = (rpi != mdt-1 ? retardo[rpi+1] : retardo[0]);
   s = retardo[rpi] + frac*(next - retardo[rpi]);
   retardo[*p] = sig[i] + s*fdb;
   sig[i] = s;
   *p = (*p != mdt-1 ? *p+1 : 0);
 return *sig;
```

Para se obter o efeito, precisamos modular o tempo de retardo com uma fonte de sinal periódica ou um envelope. No exemplo de programação abaixo, usamos um envelope simples para modificar o tempo de retardo. O código fonte completo do programa é apresentado abaixo:

```
/** flanger com envelopes.\n
flanger infile.* outfile.wav maxretardo(s) env_dur(s)
int main(int argc, char** argv) {
SNDFILE *psfo, *psfi;
int chans, bytes = sizeof(float)*def_vsize, cnt=0,
           pt=0, ts=0;
float *buffer, sr, dur, dtime, *comp, *del,
                    del1[2]={0.f, 0.f};
if(argc == 5){
  dtime = (float) atof(argv[3]);
  dur = (float) atof(argv[4]);
  buffer = new float[def_vsize];
  comp = new float[def_vsize];
  del = new float[(int)(dtime*def_sr)];
  memset(del, 0, sizeof(float)*(int)(dtime*def_sr));
  if(!(psfi = soundin_open(argv[1], &chans, &sr))){
  printf("erro ao abrir o arquivo ou entrada\n");
  exit(-1);
  if(chans > 1 || sr != def_sr) {
```

```
printf("unsupported channels or sr\n");
    exit(-1);
  if(!(psfo = soundout_open(argv[2]))){
   printf("erro ao abrir o arquivo ou saída\n");
   exit(-1);
  do {
    cnt = soundin(psfi, buffer);
    memcpy(comp, buffer, bytes);
    flanger(buffer, line(.0001f, dur, dtime, &ts),
               0.8f,dtime, del, &pt);
   balance(buffer, comp, del1);
    soundout(psfo,buffer, cnt);
   } while(cnt);
   soundin_close(psfi);
   soundout_close(psfo);
   delete[] buffer;
   delete[] comp;
   delete[] del;
  return 0;
else {
 printf("usage: retardo infile.*"
         "outf.wav maxretardo(s) env_dur(s) \n");
 return 1;
}
```

## Considerações Finais

O estudo das técnicas de síntese e processamento desenvolvido neste capítulo não é exaustivo. Existem vários detalhes, além de técnicas alternativas, relacionados com os componentes discutidos acima que não tivemos, por questões editoriais, a oportunidade de explorar. Por exemplo, quando apresentamos o tópico sobre filtros não pudemos apresentar outros tipos de filtros além de ressonadores. No entanto, o material estudado neste texto deverá servir como base para que o leitor se aprofunde em algumas das áreas introduzidas aqui. Esperamos que a leitura deste capítulo o inspire a explorar as vastas e fascinantes áreas da programação e do processamento de sinais musicais.

## 69

#### Referências

BARREIRO, Daniel; KELLER, Damián. Composição com modelos sonoros: fundamentos e aplicações eletroacústicas. In: KELLER, Damián; BUDASZ, Rogério (orgs.), *Criação musical e tecnologias: teoria e prática interdisciplinar*, p. 97-126. Goiânia: ANPOM, 2010.

BOULANGER, Richard (org.), *The Csound book*. Cambridge, Mass.: MIT Press, 2000.

DODGE, Charles; JERSE, Thomas. *Computer music*. New York: Schirmer, 1985.

LAZZARINI, Victor; TIMONEY, Joseph. New methods of formant analysis-synthesis for musical applications. *Proceedings of the International Computer Music Conference 2009*, Montreal, McGill University, 2009.

MOORE, F. Richard. *Elements of computer music*. Englewood Cliffs, NJ: Prentice-Hall. 1990.

ROADS, Curtis. *Computer music tutorial*. Cambridge, Mass.: MIT Press, 1996.

SMITH, Julius; ANGELL, James. A constant-gain digital resonator tuned by a single coefficient. *Computer Music Journal*, v. 6, n. 4, (1982), p.36-40.

SMITH, Julius. *Physical audio signal processing*. Palo Alto, CA: Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, 2004. Disponível na internet: https://ccrma.stanford.edu/~jos/pasp/

STEIGLITZ, Ken. A note on constant-gain digital resonators. *Computer Music Journal*, v. 18, n. 4 (1994), p. 8-10.

STEIGLITZ, Ken. *A signal processing primer*. Menlo Park, CA: Addison-Wesley, 1995.

WANDERLEY, Marcelo. Projeto e utilização de instrumentos musicais digitais. In: KELLER, Damián; BUDASZ, Rogério (orgs.), *Criação Musical e Tecnologias: Teoria e Prática Interdisciplinar*, p. 70-96. Goiânia: ANPOM, 2010.

# Projeto e utilização de instrumentos musicais digitais

MARCELO M. WANDERLEY

este capítulo abordaremos o projeto de instrumentos musicais digitais utilizando exemplos de dispositivos desenvolvidos no laboratório *Input Devices and Music Interaction* (IDMIL) da Universidade McGill, no Canadá.

Um instrumento musical digital – IMD (do inglês, digital musical instrument) (WANDERLEY, 2006; MIRANDA e WANDERLEY, 2006) pode ser definido como sendo composto de uma interface (gestual) hardware¹ que envia sinais de controle a algoritmos de síntese sonora² (LAZZARINI, 2010) através de estratégias de mapeamento ou correspondência de parâmetros³ pré-definidas (WANDERLEY e DEPALLE, 2004). Alem de sons, outros tipos de sinais podem ser gerados/controlados por um IMD como outros tipos de mídia (e.g. vídeo), vibrações mecânicas, forças, etc.

Este fato é importante na diferenciação entre IMDs e instrumentos musicais acústicos: no segundo caso, sons são produzidos devido às características acústicas de seus elementos vibrantes. Como nos IMDs em princípio não existem elementos vibrantes, mas algoritmos em um computador, também não existe uma relação pré-definida entre causa (por exemplo, o movimento vertical de uma tecla) e efeito (o som de uma corda excitada por um martelo de piano) e esta relação tem que ser construída pelo inventor do instrumento. Desta

<sup>&</sup>lt;sup>1</sup> Ou "controlador gestual".

 $<sup>^{\</sup>rm 2}$  Normalmente, na forma de algoritmos em um computador ou em sintetizadores.

<sup>&</sup>lt;sup>3</sup> Do inglês, mapping.

forma, não há limites quanto ao tipo de relação a adotar<sup>4</sup> ou ao tipo de mídia a ser controlada.

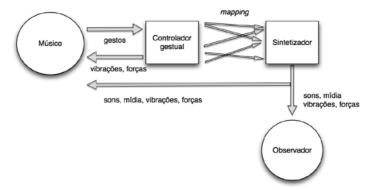


Fig 1: Uma representação de um IMD. Vê-se à esquerda a representação de um músico que interage com um controlador gestual através de gestos e movimentos, possivelmente recebendo informações deste na forma de vibrações e forças. Os dados na saída do controlador gestual são associados às entradas dos algoritmos de síntese sonora através de estratégias de correspondência ou mapeamento de dados (do inglês, mapping). O sintetizador produz sinais sonoros, mas possivelmente também visuais e vibratórios que são percebidos por observadores (e.g. plateia) e também pelo músico que interage com o IMD.

Com relação à interface de controle, esta pode, mas não precisa, ser semelhante a instrumentos acústicos existentes (JORDÀ, 2005). Com efeito, ela pode nem mesmo existir fisicamente e neste caso diríamos um "instrumento musical virtual" (do inglês, *virtual musical instrument*) (MULDER, 2000). Graças a sua inerente flexibilidade, IMDs também podem ser multiusuários, onde várias pessoas tocam um mesmo instrumento, como no caso da *ReacTable\** (JORDÀ, 2003). Também não existem mais limitações quanto ao tamanho destas interfaces ou a sua distribuição espacial. Uma interface pode estar em um mesmo local físico ou a vários quilômetros de distância, onde usuários interagem via dados enviados pela

<sup>&</sup>lt;sup>4</sup> Por exemplo, um mesmo movimento de um musico tocando um IMD pode produzir efeitos opostos.

internet (BARBOSA, 2010), como no caso da *Global String* (TANAKA E BONGERS, 2001).

Como podemos ver, praticamente não existem limitações ao projeto de interfaces, nem tampouco quanto aos sons a serem gerados, ou mesmo às estratégias de correspondência de parâmetros entre sinais de controle da interface e variáveis de entrada dos algoritmos de síntese sonora. Como então decidir quais direções tomar no projeto de novos instrumentos musicais digitais?

Neste capítulo, discutiremos o projeto e a utilização de instrumentos musicais digitais. Na primeira parte discutiremos alguns exemplos de interfaces e de instrumentos desenvolvidos no IDMIL. Na segunda parte discutiremos vários aspectos do uso destes dispositivos em diferentes contextos.

## Projeto de interfaces gestuais e instrumentos musicais digitais

Nesta seção, discutiremos alguns exemplos de interfaces gestuais clássicas e vários exemplos de dispositivos criados no laboratório IDMIL, focalizando-nos mais especificamente nas questões tecnológicas destes projetos.

#### Ponto de Partida

Como dissemos, não existem regras definidas para o projeto de interfaces de controle, para a escolha dos sons a serem controlados ou para o mapeamento entre ambos. Como então decidir qual forma estes terão?

Uma maneira de resolver este problema é a utilização de objetos do dia-a-dia associados a sensores diversos, ou partir de uma ideia composicional para tal instrumento (COOK, 2001), ou ainda de uma metáfora de controle (WESSEL e WRIGHT, 2002; FELS, GADD e MULDER, 2002).

Por exemplo, Michel Waisvisz – o finado diretor do Centro STEIM em Amsterdam e provavelmente o maior virtuoso de instrumentos digitais – queria tocar (manipular) sons com as mãos. O resultado foi a criação da interface *Hands* 

(WAISVISZ, 1985). Esta interface é composta de dois dispositivos em madeira que o intérprete utiliza como luvas em suas mãos. A distância entre estas é medida utilizando-se sensores de ultrassom e a sua orientação usando-se sensores de inclinação (tilt), além do uso de várias chaves liga-desliga para selecionar presets, por exemplo. Em várias oportunidades Waisvisz também utilizava um microfone para gravar sua voz ou ruídos da sala de concerto (por exemplo, uma salva de palmas), que eram então manipulados usando a interface Hands. Os sons utilizados, além das amostras sonoras, poderiam também incluir outros métodos de geração sonora, como a síntese por modulação de frequência.



**Fig 2**: Michel Waisvisz com a interface *Hands* em sua palestra convidada na conferência *New Interfaces for Musical Expression* (NIME03) em 24 de maio de 2003, na Universidade McGill (foto a partir do vídeo da apresentação gravado por Eileen TenCate).

É interessante observar que basicamente os mesmos gestos de Waisvisz poderiam ser capturados por outros sensores, ou por outras interfaces. Por exemplo, a distância entre as mãos poderia ser medida com sensores de luz infravermelha (IR) e a inclinação das mãos com acelerômetros. De fato, os movimentos de Alex Nowitz, que realiza performances com dois *Wiimotes* e um microfone<sup>5</sup>, muitas

<sup>&</sup>lt;sup>5</sup> Por exemplo: http://steim.org/jamboree08/?p=663 (acessado em 12 de março de 2010)

vezes lembram os movimentos nas performances de Waisvisz, mas desta vez utilizando interfaces comerciais.

Um exemplo clássico de evolução técnica de uma interface é o caso do Continuum, de Lippold Haken (HAKEN, ABDULLAH e SMART, 1992), uma interface do tamanho aproximado de um teclado de piano, mas oferecendo a possibilidade de se medir a posição de um ou mais dedos continuamente em três dimensões: X, Y e Z. A ideia por trás do conceito do *Continuum* era a de se criar uma superfície sensível ao toque para o controle de processos musicais. A primeira versão da interface usava uma câmera de vídeo para captar a posição de dedos quando estes tocavam uma superfície transparente iluminada por uma fonte de luz. As posições dos dedos criavam padrões circulares que eram identificados pela Devido a problemas de implementação tecnologia nos anos 80 (aquecimento da superfície, lentidão na análise dos padrões, sobreposição dos padrões de dedos colocados próximos), uma segunda tecnologia foi desenvolvida para implementar a mesma ideia: borracha condutora. Colocando-se fibras condutoras verticalmente em uma placa de borracha, esta pôde medir a pressão e a posição dos dedos de um músico gracas a uma matriz de contatos colocada sob a borracha. Mesmo tendo funcionado melhor do que a tecnologia anterior, ainda assim não se podia comercializar a interface, pois a placa de borracha não era suficientemente confiável, isto é, as fibras se partiam com uma utilização prolongada impedindo as medidas de pressão e de posição nas regiões onde isto acontecia. Finalmente uma terceira tecnologia foi desenvolvida, utilizando-se finas barras metálicas sob uma camada de material flexível. Estas barras, quando pressionadas, movem-se em ambas as extremidades. Sensores a efeito Hall e imãs são colocados sob cada extremidade de cada barra. Assim. quando um músico aplica uma pressão com um ou mais dedos. as barras sob os dedos deslocam-se e este deslocamento é medido pelos sensores. Note-se que, neste caso, a medida da posição no eixo horizontal (X) não é mais contínua, mas discreta (várias barras independentes). Para se obter uma medida contínua do movimento em X necessita-se interpolar os valores obtidos pelas barras próximas ao(s) ponto(s) de

75

contato. Esta terceira implementação foi finalmente comercializada por Haken<sup>6</sup> desde o fim dos anos 90.

Mas além das características puramente técnicas do projeto de interfaces e do *design* sonoro, já em 1985, Waisvisz escrevia que depois de se decidir as questões técnicas, era o momento de se aprender a tocar o novo instrumento. Quem teve a oportunidade de vê-lo ao vivo em concertos, ou mesmo por vídeo,<sup>7</sup> ficou provavelmente impressionado por sua maestria no controle de seu instrumento.

Este é um fato importante na performance com novos IMDs: por mais que se aperfeiçoem as questões técnicas, um músico terá que desenvolver uma prática instrumental para seu instrumento (BUTLER, 2008). Mas como desenvolver uma técnica instrumental para um instrumento sem história alguma e/ou para o qual não existem métodos de ensino? Discutiremos estas questões na segunda parte deste capítulo.

## Exemplos de Interfaces Gestuais Desenvolvidas no IDMIL

Nesta seção discutiremos cinco projetos de interfaces gestuais e de IMDs desenvolvidos no laboratório IDMIL da Universidade McGill: *Gyrotyre, Rulers, T-Stick, T-Box* e um teclado derivado de um piano. Estes exemplos variam de instrumentos desenvolvidos para a expressão musical a interfaces criadas para experimentos médicos em *scanners* de ressonância magnética, entre outras.<sup>8</sup>

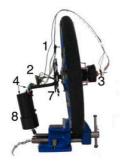
<sup>&</sup>lt;sup>6</sup> Haken Audio - http://www.cerlsoundgroup.org/Continuum/ (acessado em 15 de fevereiro de 2010)

<sup>&</sup>lt;sup>7</sup> Verificar os vários vídeos em http://www.youtube.com (procurar por "waisvisz"); acesso em 12 de março de 2010.

<sup>8</sup> Vários outros projetos são descritos em: www.idmil.org/projects/ (acessado em 12 de março de 2010)

# Gyrotyre

O *Gyrotyre* (SINYOR & WANDERLEY, 2006) é baseado em uma roda de bicicleta que pode ser manipulada segurando-a a partir de um manete acoplado a esta. Neste caso, mede-se a velocidade de rotação da roda com um giroscópio e/ou com sensores de luz infravermelha ou de efeito *Hall*, sua inclinação em dois eixos com um acelerômetro, a força com que se segura o manete com resistores sensíveis à força aplicada (do inglês, *Force Sensing Resistors* – FSRs), além de chaves liga-desliga.





**Fig 3**: o *Gyrotyre*, de Elliot Sinyor. As partes indicadas na figura são: 1) sensor a efeito *Hall*, 2) acelerômetro, 3) giroscópio, 4) resistor sensível à força, 5) imã (utilizado com o sensor a efeito *Hall*), 6) fotodiodo, 7) suporte circular para os fotodiodos(um diodo emissor de luz é colocado na roda) e 8) manete (SINYOR e WANDERLEY, 2006).

O uso de múltiplos sensores para medir a velocidade de rotação da roda de bicicleta se explica pelas limitações técnicas de cada tipo de sensor. Por exemplo, o giroscópio usado mede ao máximo rotações de 300/s, quer dizer, menos de uma volta por segundo. Por outro lado, ambos os sensores de infravermelho e de efeito *Hall* podem medir vários pulsos a

77

cada volta, dependendo do número de emissores e de receptores colocados sobre a roda<sup>9</sup>.

A ideia por trás desta interface era a de se utilizar sistemas mecânicos com propriedades dinâmicas próprias (SINYOR, 2006). Como a roda gira em torno de um eixo, um movimento perpendicular a este eixo vai causar uma resposta equivalente no terceiro eixo (precessão), perpendicular a estes dois, devido à lei de conservação do momento angular (FRADEN, 2004). Este fenômeno está na base do desenvolvimento de giroscópios mecânicos no século XIX, daí o nome escolhido para interface, *Gyrotyre*.

Como a interface tem um comportamento dinâmico próprio, 10 o usuário deve levar em conta este comportamento quando tocar um IMD que a utiliza. De modo similar ao que acontece com instrumentos acústicos quando se muda de registro, por exemplo, em uma clarineta, o músico deve adaptar seus gestos para manter uma continuidade sonora (nível sonoro, frequência e timbre) entre registros. O mesmo se aplica ao *Gyrotyre*: movendo-se a interface em diferentes direções causa efeitos distintos, forçando o músico a adaptar seu gesto ao comportamento da interface. O objetivo final é obter-se mais expressividade na performance graças a este comportamento dinâmico e ao aprendizado do músico para controlá-lo.

<sup>&</sup>lt;sup>9</sup> Por exemplo, se colocarmos um diodo emissor de luz infravermelha (*LED* IR) e um fotodiodo receptor de IR, a cada volta completa da roda teremos um pulso quando ambos se posicionam face à face. Se por acaso utilizarmos múltiplos emissores (ou múltiplos receptores), por exemplo 2 emissores, teremos múltiplos pulsos (i.e., 2 pulsos) por volta, e assim por diante. O mesmo é válido para o caso de sensores a efeito *Hall*, onde se utilizariam sensores e imãs. Em ambos os casos, esta solução é interessante para velocidades angulares elevadas, i.e., várias voltas por segundo. Para se medir baixas rotações, o sinal contínuo do giroscópio é mais indicado.

<sup>10</sup> Isto é, não é completamente passiva, mas também não utiliza motores para gerar energia mecânica.

#### Rulers

A interface *Rulers* é composta de várias "réguas" ou hastes metálicas de diferentes comprimentos e foi desenvolvida por David Birnbaum (MALLOCH, BIRNBAUM, SINYOR e WANDERLEY, 2006). Abaixando-se ou levantando-se estas hastes, assim como percutindo-as, suas vibrações e diferentes posições são captadas por sensores infra-vermelhos colocados abaixo delas.

Deste modo, pode-se tocá-la de um modo similar a um carrilhão de igreja, ou seja, percutindo as hastes, ou modulando continuamente para cima ou para baixo suas posições quando as pontas destas são seguradas com os dedos, ou mesmo com os braços.





**Fig 4**: Esquerda: *T-Box*, de Geoff Holbrook, Eileen TenCate e Mark Marshall. Os quatro receptores de ultrassom são colocados na caixa de madeira. Os dois emissores estão colocados nas tiras em cada mão. Direita: *Rulers*, de David Birnbaum. Cada haste metálica (alumínio) oscila de maneira diferente devido aos vários comprimentos (fotos: Vanessa Yaremchuk, músico: Fernando Falci de Souza)

#### T-Stick

A ideia por trás do *T-Stick*, desenvolvido por Joseph Malloch (MALLOCH e WANDERLEY, 2006), era a de se construir uma interface em forma de bastão que fosse sensível ao toque

Mark to the control of the control o

em toda a sua superfície. Alguns dos principais requisitos do

projeto eram:

- Metáfora de controle: interação com uma corda ou barra virtual, com a possibilidade de se excitar ou atenuar a vibração destas estruturas;
- Sensores e mapeamento deveriam ser definidos de forma integral, isto é, capazes de medir gestos de maneira simultânea e de forma inter-relacionada;
- Simplicidade de operação: a interface deveria ser simples de ser operada, tal como um instrumento elétrico;
- Robustez: a interface deveria poder ser tocada durante várias horas por dia sem que esta parasse de funcionar.

O primeiro *T-Stick* foi construído a partir de um tubo plástico de 120 cm, dividido em dois, de forma a permitir que dispositivos eletrônicos fossem colocados dentro do tubo.



**Fig 5**: Vista parcial dos circuitos e conexões dentro de um *T-Stick*. (Foto de Joseph Malloch)

Vários sensores capacitivos (campo elétrico) discretos (liga-desliga) são utilizados para medir se o músico toca um ponto determinado da interface, sendo o número final de sensores dependente da resolução esperada. Tiras de cobre são usadas como eletrodos destes sensores e cobrem a superfície de um dos lados do tubo.



Fig. 6: O primeiro *T-Stick* construído. (Foto de Joseph Malloch)

Outros sensores utilizados são dois acelerômetros (de 3 eixos cada) colocados em cada extremidade do tubo, sensores de pressão feitos com papel condutor (KOEHLY, CURTIL e WANDERLEY, 2006) na metade do tubo onde não são colocados sensores capacitivos, e um cristal piezoelétrico colado ao tubo para medir tensões na estrutura resultantes de torções e impactos.

Uma outra característica importante do projeto de desenvolvimento do *T-Stick* era a utilização da ideia de "famílias de instrumentos" <sup>11</sup> no contexto de interfaces gestuais de IMDs (MALLOCH, 2007). No caso do *T-Stick*, foram criados diversos modelos da interface de controle com diferentes dimensões (*T-Stick* tenor e *T-Stick* soprano), resoluções (24 ou 48 eletrodos para sensores capacitivos em uma mesma distância), saídas de dados (cabo USB ou sem fio), ou variações em alguns dos tipos de sensores utilizados. Por exemplo, um dos *T-Stick* recentemente construído utiliza, além dos sensores descritos acima, sensores de luz ambiente (resistores sensíveis à luz), sensores de sopro e sensores de proximidade utilizando luz infravermelha.

<sup>&</sup>lt;sup>11</sup> Por exemplo, a família das cordas, composta pelos violinos, violas, violoncelos e contrabaixos. Estes vários instrumentos musicais têm características comuns (o uso de cordas que serão excitadas por um arco ou pelos dedos) e são tocados de modo similar.

81

Uma nova geração de *T-Sticks* utiliza sensores inerciais e de posição (giroscópio e compasso eletrônico), além dos acelerômetros, para indicar a orientação espacial da interface em tempo real. Este *T-Stick* foi construído para um projeto<sup>12</sup> envolvendo um dançarino que utilizava esta interface enquanto dançava para *espacializar* sons gerados por um violoncelo em um ou mais de 16 alto-falantes distribuídos no espaço da performance.

Independentemente do tipo de *T-Stick*, todas as interfaces gestuais compartilham as mesmas características básicas descritas acima e podem assim ser tocadas de forma similar. Obviamente, a questão dos timbres produzidos por uma mesma família de instrumentos acústicos não se aplica ao caso dos IMDs.

#### T-Box

A história da *T-Box* começou com o compositor Geoff Holbrook. A interface original utiliza sensores de ultrasom para medir a distância entre as mãos de um músico (MIRANDA & WANDERLEY, 2006, pg. 128) de maneira similar as *Hands* de Michel Waisvisz. Holbrook usou a medida da intensidade do sinal de ultrasom emitido (40kHz) por um canal de uma placa de som, contrariamente à medida do tempo de propagação de um pulso (em inglês, time of flight) que é uma técnica mais comum. A vantagem da medida de intensidade é sua simplicidade, isto é, envia-se um sinal com intensidade conhecida e mede-se este sinal na chegada. Entretanto, suas limitações incluem a não linearidade da medida assim como o fato de que a intensidade do sinal diminui tanto com um aumento da distancia emissor-receptor quanto com uma variação no ângulo de um deles com relação à linha imaginária que os une.

A versão utilizada no projeto Orquestra Digital<sup>13</sup> foi alterada por Holbrook e por Eileen TenCate para utilizar uma

<sup>&</sup>lt;sup>12</sup> Coordenado pelo compositor Sean Ferguson e pela coreógrafa Isabelle Van Grimde.

<sup>13</sup> Ver descrição deste projeto na segunda parte deste artigo.

referência fixa onde os receptores são colocados. Mede-se assim a distância entre os (2) emissores, um em cada mão e os (4) receptores na referência. 14 O fato de medir tanto a distância quanto o ângulo permite o desenvolvimento de técnicas de performance onde o músico é capaz de controlar variáveis continuamente (usando principalmente a distância) ou de maneira discreta (usando principalmente o ângulo), criando assim mais condições para aumentar a expressividade no uso desta interface.

## Teclado compatível com scanners de ressonância magnética

O projeto de interfaces musicais para o uso em situações medicais traz novos desafios para o projetista. Neste caso, fomos contatados pelo neurologista Robert Zatorre (Instituto Neurológico de Montreal, Universidade McGill) e pela psicóloga Virginia Penhune (Universidade Concórdia) para saber se seria possível desenvolver instrumentos musicais digitais compatíveis com exames de ressonância magnética.

Um scanner de ressonância magnética (SRM) é um aparelho capaz de detectar fluxo sanguíneo no cérebro usando campos magnéticos estáticos e pulsantes. Desta forma, dispositivos que contenham materiais metálicos, em especial materiais ferromagnéticos, não podem ser utilizados com estes sistemas, pois seriam atraídos com violência para o centro do scanner. Assim, a maioria dos instrumentos musicais existentes não podem ser usados neste contexto, não só pela presenca de metais, mas também devido às reduzidas dimensões dos scanners e aos ruídos produzidos em funcionamento. Desta forma, se guisermos verificar o que acontece no cérebro de um músico enquanto este interpreta uma partitura no scanner, devemos desenvolver instrumentos compatíveis com estes Na impossibilidade de se desenvolver sistemas. instrumentos, muitas das pesquisas atuais com músicos em SRM são feitas com teclados de plástico onde se remove as

 $<sup>^{14}</sup>$  Por exemplo, em duas hastes de madeira em forma de cruz ou , em uma versão mais recente desenvolvida por Mark Marshall, em um caixa.

83

partes eletrônicas, ou mesmo com os músicos fingindo tocar um teclado, mas com os dedos no ar.

Avrum Hollinger desenvolveu no IDMIL um teclado que utiliza fibras óticas e a medida de intensidade luminosa para se determinar a posição e a velocidade de 11 teclas de piano <sup>15</sup> (HOLLINGER, PENHUNE, ZATORRE, STEELE e WANDERLEY, 2007). Com a escolha de fibras óticas, pode-se colocar os circuitos de controle e de condicionamento das medidas a vários metros de distância da interface em uma sala separada, e assim evitar possíveis problemas de segurança.

As principais características de uma interface musical compatível com *scanners* de ressonância magnética são:

- Segurança para os utilizadores;
- Não causar interferência nas medições obtidas pelo *scanner*;
- Confiabilidade.

A primeira característica diz respeito à segurança das pessoas envolvidas nos testes. É óbvio que não se pode tolerar quaisquer riscos à saúde de músicos, técnicos ou pesquisadores enquanto a interface é utilizada na sala de exames. Mas além das questões óbvias de segurança, também é importante que a interface não interfira no funcionamento do *scanner*, isto é, que não haja distorções nas medidas obtidas.

Finalmente, é essencial que a interface seja confiável. A necessidade de uma interface confiável vem da necessidade que os dados obtidos sejam os mais exatos (isto é, pequenos erros) e precisos (isto é, boa repetição das medidas) possível. Também é importante desenvolver uma interface de fácil utilização para técnicos de serviços de medida por ressonância magnética, de forma que os projetistas não necessitem estar constantemente à disposição durante exames.

 $<sup>^{15}</sup>$  O tamanho escolhido para a escala da interface (menos de 1 oitava) é principalmente função das limitações de espaço no scanner.





**Fig 7**: O protótipo de interface de piano desenvolvido no IDMIL por Avrum Hollinger e um exemplo de sua utilização durante um teste em um SRM no Instituto Neurológico de Montreal. (Fotos: Avrum Hollinger)

No caso do teclado compatível com SRM, vemos que não somente a escolha de sensores, mas também o tamanho e a funcionalidade da interface gestual, foram feitas em função da aplicação e não unicamente das características desejadas pelo projetista ou pelo músico que a utilizará ou mesmo do objetivo artístico desta utilização.

# Utilização de instrumentos musicais digitais

Na segunda parte deste capítulo discutiremos a utilização de interfaces e IMDs em vários contextos musicais, assim como a influência destes no projeto de novos instrumentos. Levantaremos várias questões ligadas à performance de IMDs e proporemos algumas respostas baseadas na experiência obtida durante o projeto Orquestra Digital.

Uso de instrumentos musicais digitais em concertos e performances

Como descrito no caso das *Hands* de Michel Waisvisz, talvez o mais óbvio uso de IMD é como uma ferramenta de expressão musical em concertos e performances em geral. Nestas condições, um IMD pode ser utilizado de forma similar a um instrumento acústico, isto é, em interpretações de peças pré-compostas, em improvisações, em solo ou em grupos.

Obviamente, IMDs podem ser usados em outras aplicações, por exemplo, como ferramentas pedagógicas. Na Figura 8 pode-se ver crianças interagindo com um *T-Stick* durante o festival *Wired NextFest* 2007 em Los Angeles, Califórnia.



**Fig 8**: Três crianças brincando com um *T-Stick* durante o festival *Wired NextFest* 2007. (Foto: Joseph Malloch)

Nesta situação, não existe a possibilidade de aprendizado de técnicas de interpretação por um músico que desenvolve gradualmente seu controle do instrumento. De fato, já que não se pode prever exatamente o que vai acontecer, o funcionamento do instrumento deve ser suficientemente claro para que as crianças se interessem por ele, mas também suficientemente robusto para suportar ações inesperadas de dezenas ou centenas de crianças.

No caso do teclado compatível com *scanners* de ressonância magnética, vemos que, diferentemente da utilização em concertos ou eventos musicais, as maiores implicações no projeto são de ordem do ambiente em que esta interface é utilizada.

# Protótipos ou Instrumentos

Independente da utilização almejada, é imprescindível que um IMD se comporte de maneira previsível e estável.

Embora estas características possam ser consideradas óbvias em instrumentos acústicos (mas nem sempre), uma grande parte das novas interfaces propostas não se comportam desta forma.

De fato, um dos mais importantes passos na criação de um IMD é a passagem de um protótipo de laboratório a um instrumento que pode ser utilizado profissionalmente em vários ambientes externos ao laboratório onde este foi concebido. Esta questão é de vital importância em laboratórios têm-se normalmente ambientes controlados (luz, temperatura, ruído, etc.) O uso de um IMD fora do laboratório vai expô-lo a propriedades de ambientes que podem afetar seu funcionamento.

Um exemplo clássico é o uso de sensores de luz IR para medir posição ou a vibração de uma corda (OVERHOLT, 2005). Esta técnica de medição em geral funciona bem e é de baixo custo. Mas quando a se utiliza em concertos, luzes em um palco ou flashes de câmeras fotográficas podem causar interferências e erros nas medidas dos sensores, erros que não serão detectados em um laboratório com um ambiente controlado. A interface de controle neste caso pode não funcionar ou funcionar de maneira defeituosa.

O mesmo pode ocorrer com vários outros tipos de sensores, por exemplo, sensores magnéticos, quando a presença de metais em um ambiente pode modificar seu funcionamento, ou com sensores de ultrasom ou extensômetros (do inglês, *strain gauges*), que são sensíveis à variações de temperatura e/ou de umidade.

A passagem de um protótipo que funciona bem em um laboratório a um instrumento que pode ser usado em diversas situações foi um grande desafio no projeto Orquestra Digital, cujos IMDs foram desenvolvidos no IDMIL.

A ideia por trás da Orquestra Digital era de se utilizar IMDs em um contexto de performance musical onde peças seriam compostas principalmente por estudantes de composição e executadas por estudantes de mestrado e doutorado em performance da universidade McGill. Neste caso, os vários desenvolvimentos do IDMIL tiveram que ser

aprimorados para poder ser utilizados em concertos públicos 16. Estes deveriam se comportar como instrumentos acústicos, de modo que um músico competente pudesse utilizá-los sem a ajuda de um técnico em permanência a seu lado.



**Fig 9:** Vários instrumentos, músicos e engenheiros do projeto Orquestra Digital. Da esquerda para a direita: o engenheiro Mark Marshall, utilizando os emissores da interface *T-Box*; a violoncelista Cloé Domingues, com luvas aumentadas com sensores de pressão isométricos (resistores sensíveis à força) desenvolvidas por Pierre-Yves Fortier e por Mark Marshall; Fernando Rocha, professor de percussão na UFMG e na época estudante de doutorado em interpretação na McGill, com o *T-Stick*; o compositor D. Andrew Stewart com as *Rulers*. (Foto: Joseph Malloch).

Mas não só as interfaces de controle necessitaram aprimoramentos. Como se utilizam computadores de uso geral para gerar os sons controlados pelas interfaces, e como estes computadores não são nem projetados, nem adaptados para usos musicais onde restrições de tempo real são importantes<sup>17</sup>

 $<sup>^{16}</sup>$  Que aconteceram em Março de 2008 na sala Pollack da universidade McGill.

<sup>&</sup>lt;sup>17</sup> Um projeto recente no IDMIL utiliza uma plataforma hardware baseada no microcontrolador ARM7 para gerar sons, de forma a

se comparadas ao seu uso em atividades de escritório (processamento de textos) ou de navegação de *websites*, o desenvolvimento e a utilização de softwares para o mapeamento e para a síntese sonora tiveram que levar em conta as necessidades de uso em concertos (MALLOCH, SINCLAIR e WANDERLEY, 2008).

### Flexibilidade ou limitações

Quando se inicia o projeto de um IMD, uma ideia frequente é a de se fazer instrumentos flexíveis ou adaptáveis ao músico e/ou ao ambiente. Desta forma, tenta-se incorporar um grande número de características a um IMD, principalmente em nível de programação software. Mas seria a flexibilidade total uma necessidade?

Novamente, se analisarmos instrumentos musicais acústicos, veremos que algumas das técnicas de interpretação foram construídas em torno de limitações destes objetos. Um exemplo são as técnicas para se tocar sons multifônicos em instrumentos de sopro, que basicamente tentam obter sons estáveis em uma região de funcionamento instável do instrumento. Esta discussão se baseia na pergunta: Um instrumento que não tenha limites pode ser tocado? Obviamente, a definição da flexibilidade ou dos limites de um IMD deve ser função dos objetivos do projetista, mas pessoalmente acho importante que IMDs tenham um caráter próprio, incluindo limitações que serão utilizadas por músicos quando estes os utilizarem em suas interpretações.

# Ergonomia

A questão da ergonomia de IMDs também é um fator muito importante. Não se procura necessariamente fazer instrumentos fáceis de tocar – ver por exemplo a discussão da relação entre esforço e expressão musical, por Joel Ryan (1992) – mas aprendemos durante o projeto Orquestra Digital que

possibilitar a geração de sons junto à interface sem a necessidade de utilizar computadores para este fim (HOLLINGER, THIBODEAU e WANDERLEY, 2010).

89

músicos não vão necessariamente aceitar aprender a tocar a interface de um instrumento que requeira esforço físico suscetível de causar problemas físicos. Este foi o caso com o *Gyrotyre*.



**Fig 10**: Elliot Sinyor demonstrando o *Gyrotyre*. (Fotos a partir de um vídeo feito por Elliot Sinyor)

Pessoalmente acho o *Gyrotyre* uma interface bastante interessante e com um alto potencial expressivo<sup>18</sup>, mesmo assim este não foi escolhido pelos músicos participantes no projeto como parte dos instrumentos a serem utilizados no concerto final. A razão principal foi o seu peso e o esforço do antebraço, necessário para suportá-lo no ar. Como os músicos do projeto tocavam profissionalmente instrumentos acústicos, estes preferiram não se lançar no aprendizado do *Gyrotyre* por medo de sofrerem lesões, e com toda a razão, pois uma possível lesão causaria problemas em suas carreiras profissionais.

 $<sup>^{\</sup>rm 18}$  Vídeos em www.youtube.com/idmil (acessado em 12 de março de 2010)

## Composição - Notação

Como dissemos, os sons gerados por um IMD são definidos pelo inventor da interface e/ou por seus usuários. Qual é então a "voz" de um instrumento musical digital?<sup>19</sup> Esta questão tem uma consequência direta nas estratégias de composição musical utilizando IMDs: Como compor para instrumentos que não têm tessituras pré-definidas? No caso da Orquestra Digital, como várias peças foram compostas para mais de um IMD, foi necessário definir características sonoras para cada um deles<sup>20</sup> (FERGUSON e WANDERLEY, 2009).

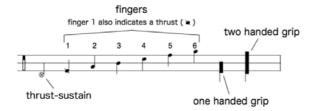
Outra questão importante é a escrita de partituras para IMDs. Como escrever a partitura de uma peça composta, com símbolos musicais que indicam a altura e a intensidade de uma nota como para instrumentos acústicos, ou anotando-se os gestos necessários para tocar tal instrumento? Talvez uma combinação de ambos seja uma solução, pois como não existem métodos para novas interfaces, não existem informações de como tocá-las no contexto de uma peça.

Na figura 11 vemos a solução adotada por D. Andrew Stewart em sua peça *Catching Air and the Superman* (STEWART, 2010), para 2 *T-Sticks* sopranos, teclado eletrônico, e orquestra de câmara. A parte do *T-Stick* é anotada em três linhas, sendo que a superior e a inferior representam os extremos superior e inferior da superfície coberta pelos sensores capacitivos. Vê-se assim as indicações dos gestos necessários para se produzir os sons desejados.

\_

<sup>&</sup>lt;sup>19</sup> Outra questão importante é como combinar sons sintetizados (e.g. sons eletrônicos criados por criados por IMDs) com sons de instrumentos acústicos em peças mistas (SOUZA, 2010).

<sup>&</sup>lt;sup>20</sup> Estas características podem obviamente variar entre aplicações.



**Fig 11:** Um exemplo da notação desenvolvida para o *T-Stick*, pelo compositor D. Andrew Stewart em sua peça *Catching Air and the Superman* (2010). Vê-se a notação do dedilhado indicando o dedo a utilizar e o seu posicionamento no instrumento. Também se vê uma maneira de usar o *T-Stick*: com um movimento rápido sacudindo o instrumento na direção de seu eixo (em inglês, *thrust*) ou aplicando uma pressão contínua com as mãos antes de fazer este movimento (em inglês, *thrust-sustain*). Finalmente, tem-se a indicação de como segurar a interface, com uma mão (em inglês, *one handed grip*) ou com as duas mãos (em inglês, *two handed grip*)

# Interpretação

Mas uma questão ainda não foi abordada diretamente: Como desenvolver uma técnica instrumental para uma nova interface gestual ou IMD? Isto é, como desenvolver e depois ensinar a maneira de tocar este novo instrumento? Como já dissemos, como não existem limitações acústicas (por exemplo, a vibração de uma membrana), a flexibilidade no projeto é total. Poderemos sempre nos referir às possíveis *affordances* da interface, isto é, quais movimentos seriam os mais óbvios ou naturais quando as utilizamos, mas mesmo assim não existe uma resposta única a esta questão.

Novamente, se nos referirmos ao caso do *T-Stick*, pelo menos três técnicas de interpretação foram desenvolvidas para esta interface, como ilustrado na figura 12. Note-se que, na figura da esquerda, o *T-Stick* tem uma haste de metal acoplada a uma de suas extremidades para mantê-lo no solo, de maneira semelhante a um violoncelo. A técnica de performance principal neste caso consiste em aplicar torções em torno de seu eixo, assim como deslizar a mão direita na direção do eixo da interface para se criar sons ou para modular sons existentes. Na foto central o *T-Stick* é mantido no ar. Neste caso, a técnica

principal consiste em dedilhados acionando os sensores capacitivos (geração de notas musicais) e na constante inclinação do instrumento. Finalmente na foto da direita a técnica consiste em movimentos bruscos, que também incluem variações na forma como se segura a interface, mas não inclui dedilhados (fotos feitas a partir dos vídeos dos concertos).

As duas primeiras técnicas foram desenvolvidas simultaneamente pela pianista Xenia Pestova e pelo percussionista Fernando Rocha, ambas durante um curso sobre IMDs na universidade McGill em 2006, quando o *T-Stick* foi concebido. Na parte direita da figura 12 vê-se uma terceira técnica de interpretação do *T-Stick*, desta vez desenvolvida pelo compositor D. Andrew Stewart. Neste caso, Stewart usa a interface como um bastão de artes marciais (sem necessariamente batê-lo contra uma superfície), utilizando gestos mais abruptos. Note-se que, neste caso, ele toca um *T-Stick* soprano, com a metade do tamanho de um *T-Stick* tenor.







**Fig 12**: Xenia Pestova (esquerda) e Fernando Rocha (centro) com o primeiro protótipo de um *T-Stick* (tenor) e D. Anfrew Stewart (direita) interpretando sua peça "Everybody to the Power of One", para *T-Stick* soprano.

Entre estas três opções, qual seria então a técnica mais apropriada ao *T-Stick*? Todas ou talvez nenhuma, a resposta depende do intérprete que utilizará a interface. Mas no caso do T-Stick tivemos a sorte de contar com dois intérpretes (Rocha e Stewart) que se dedicaram durante anos a inventar e a

aprimorar suas técnicas instrumentais e que continuam a desenvolver suas habilidades com este instrumento, tornandose ambos virtuosos do *T-Stick*. Talvez outros músicos se inspirarão destes exemplos para começar carreiras como intérpretes de instrumentos musicais digitais e criarão eles mesmos novas técnicas de interpretação.

#### Conclusões

Neste capítulo discutimos exemplos de projetos de interfaces gestuais de instrumentos musicais digitais, assim como suas utilizações em diferentes contextos. Depois de passar em revista alguns dos dispositivos criados laboratório IDMIL da Universidade McGill, discutimos seu uso em projetos artísticos como a Orquestra Digital ou em exames médicos para a pesquisa sobre a performance musical utilizando scanners de ressonância magnética. Vimos que estes contextos definem em grande parte a forma final de um IMD, e que múltiplas soluções existem para o projeto e para a interpretação musical com estes novos instrumentos. Como não existem soluções únicas (nem necessariamente óbvias), é necessário muito trabalho e dedicação para desenvolver novos instrumentos e técnicas de interpretação que possibilitem a músicos explorar de forma satisfatória as possibilidades musicais oferecidas pela tecnologia digital.

# Agradecimentos

Gostaria de agradecer aos alunos e aos pesquisadores citados neste capítulo pela colaboração durante vários anos no IDMIL e pelo uso de imagens de seus trabalhos principalmente: David Birnbaum, Avrum Hollinger, Joseph Malloch, Mark Marshall, Fernando Rocha, Elliot Sinyor e D. Andrew Stewart. Também gostaria de agradecer às agencias de financiamento de pesquisa Natural Sciences and Engineering Council of Canada (NSERC) e Canadian Foundation for Innovation (CFI), pelo suporte a esta linha de pesquisa e ao Fonds Québécois de Recherche sur la Société et la Culture (FQRSC), pelo financiamento do projeto Orquestra Digital. Obrigado

finalmente a Fernando Falci de Souza e aos revisores anônimos pela releitura atenta deste capítulo.

#### Referências

BARBOSA, Álvaro. Performance musical em rede. In: KELLER, Damián; BUDASZ, Rogério (orgs.), *Criação musical e tecnologias: teoria e prática interdisciplinar*, p. 178-199. Goiânia: ANPOM, 2010.

BUTLER, Jennifer. Creating pedagogical etudes for interactive instruments. In: *Proceedings of NIME08* – International Conference on New Interfaces for Musical Expression, Genoa, 2008, p. 77-80, 2008.

COOK, Perry. Principles for designing computer music controllers. In: *Proceedings of NIME01*- New Interfaces for Musical Expression Workshop, Seattle, 2001. Disponível na internet: http://www.nime.org/2001/ (acessado em 12 de março de 2010)

FRADEN, Joseph. *Handbook of modern sensors: physics, design and applications.* London: Springer, 2004.

FELS, Sidney; GADD, Ashley; MULDER, Axel. Mapping transparency through metaphor: toward more expressive musical instruments, *Organized Sound*, v. 7, n. 2 (2002), p. 109–126.

FERGUSON, Sean; WANDERLEY, Marcelo. The McGill Digital Orchestra: interdisciplinarity in digital musical instrument design. In: *Proceedings of the 5th International Conference on Interdisciplinary Musicology*, Paris, 2009.

HAKEN, Lippold; ABDULLAH, Radi; SMART, Mark. The continuum: a continuous music keyboard. In: *Proceedings of ICMC92* – International Computer Music Conference, p. 81–84, 1992.

HOLLINGER, Avrum; PENHUNE, Virginia; ZATORRE, Robert; STEELE, Chros; WANDERLEY, Marcelo. fMRI-Compatible electronic controllers. In: *Proceedings of NIME07* – International Conference on New Interfaces for Musical Expression, Nova York, p. 246-249, 2007.

HOLLINGER, Avrum; THIBODEAU, Joseph; WANDERLEY, Marcelo. An embedded hardware platform for fungible interfaces. *Proceedings of ICMC10* – International Computer Music Conference, 2010. (No prelo).

JORDÅ, Sergi. Sonigraphical instruments: from FMOL to the reacTable\*. In: *Proceedings of NIME03* – International Conference on New Interfaces for Musical Expression, Montreal, Canada, p. 70-76, 2003.

JORDÀ, Sergi. *Digital Lutherie: Crafting musical computers for new musics' performance and improvisation*. Barcelona, 2005. Tese de doutorado – Universitat Pompeu Fabra.

KOEHLY, Rodolphe; CURTIL, Denis; WANDERLEY, Marcelo. Paper FSRs and latex/fabric traction sensors: methods for the development of home-made touch sensors. In: *Proceedings of NIME06* – International Conference on New Interfaces for Musical Expression, Paris, p. 230-233, 2006.

LAZZARINI, Victor. Técnicas fundamentais para instrumentos computacionais de síntese e processamento de áudio. KELLER, Damián; BUDASZ, Rogério (orgs.), *Criação musical e tecnologias: teoria e prática interdisciplinar*, p. 39-69. Goiânia: ANPOM, 2010.

MALLOCH, Joseph; WANDERLEY, Marcelo. The T-Stick: from musical interface to musical instrument. In: *Proceedings of NIME07 – International Conference on New Interfaces for Musical Expression*, Nova York, p. 66-69, 2007.

MALLOCH, Joseph; BIRNBAUM, David; SINYOR, Elliot; WANDERLEY, Marcelo. A New Conceptual Framework for Digital Musical Instruments. In: *Proceedings of the DAFx-06* – 9th International Conference on Digital Audio Effects, Montreal, p. 49-52, 2006.

MALLOCH, Joseph. *A consort of gestural musical controllers: design, construction, and performance.* Montreal, 2007. Dissertação de Mestrado – McGill University.

MALLOCH, Joseph; SINCLAIR, Stephen; WANDERLEY, Marcelo. A network-based framework for collaborative development and performance of digital musical instruments. In: KRONLAND-MARTINET, Richard; YSTAD, Sølvi; JENSEN, Kristoffer (orgs.), Computer music modeling and retrieval: sense of sounds. Proceedings of the 4th Computer Music Modeling and Retrieval Conference, p. 401–425. Berlim / Heidelberg: Springer-Verlag, 2008. Série "Lecture Notes in Computer Science", v. 4969.

MIRANDA, Eduardo; WANDERLEY, Marcelo. *New digital musical instruments: control and interaction beyond the keyboard.* Madison: A-R Editions, 2006.

MULDER, Axel. Toward a choice of gestural constraints for instrumental performers. In: WANDERLEY, Marcelo; BATTIER, Marc (orgs.), *Trends in gestural control of music*, p. 315-335. Paris: Institut de Recherche et Coordination Acoustique Musique, 2000.

OVERHOLT, Dan. The overtone violin. In: *Proceedings of NIME05* – International Conference on New Interfaces for Musical Expression, Vancouver, p. 34-37, 2005.

RYAN, Joel. Effort and Expression. In: *Proceedings of ICMC92* – International Computer Music Conference, p. 414–416, 1992.

SINYOR, Elliot. "Digital Musical Instruments: A Design Approach Based on Moving Mechanical Systems", *M.A. thesis*, McGill University. August 2006

SINYOR, Elliot; WANDERLEY, Marcelo. Gyrotyre: a hand-held dynamic computer-music controller based on a spinning wheel. In: *Proceedings of NIME05 – International Conference on New Interfaces for Musical Expression*, Vancouver, p. 42-45, 2005. Stewart, 2010.

SOUZA, Rodolfo Coelho de. Da interação entre sons instrumentais e eletrônicos. In: KELLER, Damián; BUDASZ, Rogério (orgs.), *Criação musical e tecnologias: teoria e prática interdisciplinar*, p. 148-177, Goiânia: ANPOM, 2010.

STEWART, David Andrew. *Catching air and the Superman*. Montreal, 2010. Tese de doutorado – McGill University.

TANAKA, Atau; BONGERS, Bert. Global string: a musical instrument for hybrid space. In: *Proceedings of Cast01: Living in Mixed Realities*, Sankt Augustin, Alemanha, p. 177–81, 2001.

WAISVISZ, Michel. The Hands: A set of remote MIDI-controllers. In: *Proceedings of ICMC85 – International Computer Music Conference*, p. 313–18, 1985.

WANDERLEY, Marcelo. Instrumentos musicais digitais: gestos, sensores e interfaces. In: ILARI, Beatriz (org.), *Em busca da mente musical*, p. 70-96. Curitiba: Editora da Universidade Federal do Paraná, 2006.

WANDERLEY, Marcelo; DEPALLE, Philippe. Gestural control of sound synthesis. *Proceedings of the IEEE*, v. 92, n. 4 (abr. 2004), p. 632-644.

WESSEL, David; WRIGHT, Matthew, Problems and Prospects for Intimate Musical Control of Computers, *Computer Music Journal*, v. 26, n. 3 (2002), p. 11–22.

# Composição com modelos sonoros: fundamentos e aplicações eletroacústicas

DANIEL LUÍS BARREIRO

DAMIÁN KELLER

I

## A síntese sonora no contexto da prática musical

Prês aspectos da atividade composicional com sons estão intimamente entrelaçados: a síntese sonora, a geração de parâmetros e a exploração das possibilidades fornecidas pelas técnicas adotadas. Tendo em vista a interação entre esses três aspectos, discutimos o impacto das atividades musicais no desenvolvimento das ferramentas composicionais. Dependendo das variáveis do processo de geração sonora, o músico pode ter acesso de forma direta e intuitiva a parâmetros sonoros consistentes com o enfoque composicional adotado, ou pode precisar modificar a lógica de funcionamento dos sistemas de síntese e processamento para atingir seus objetivos musicais. Um conceito chave no desenvolvimento de ferramentas de controle intuitivas é o canal de interação ou affordance (KELLER et al., 2009). A presente seção introduz o conceito de canal de interação dentro do contexto das atividades musicais. A partir da descrição dos níveis de organização sonora, estabelecemos um marco teórico que nos permitirá abordar o estudo dos parâmetros de síntese e processamento nos contextos composicionais focalizados na seção seguinte.

#### O fazer musical como atividade

Se entendemos a música como um conjunto de atividades, a criação musical torna-se mais do que a escrita de símbolos em uma folha pautada. Fazer música implica produzir som – e não apenas no nível da representação de ações para uma fonte sonora individual, como é o caso da escrita instrumental em notação tradicional, mas também manipulação de parâmetros nos níveis micro, meso, macro e no contexto extramusical (KELLER e CAPASSO, 2006). A atividade musical não se limita ao trabalho isolado do músicocompositor, ela demanda um processo de compartilhamento social indispensável para o surgimento e a consolidação das novas práticas musicais. Portanto, a criação musical envolve pelo menos três tipos de atividades: 1) As atividades individuais que não têm um resultado sonoro direto; 2) As atividades individuais com o objetivo imediato de produzir som; e 3) A interação social. O terceiro item é altamente relevante no contexto de sistemas que envolvem a participação ativa de usuários múltiplos. No presente artigo estaremos focalizando unicamente sistemas em que a composicional é individual, portanto indicamos a leitura de Keller et al. (2009), Pimenta et al. (2009), e Miletto et al. (2009) para mais detalhes sobre o conceito de interação social no contexto composicional.

Em Keller *et al.* (2009) introduzimos o uso da teoria da atividade no contexto musical (LEONT'EV, 1978). A unidade de análise no estudo da atividade humana é o sistema de atividades, isto é, uma comunidade de atores ou agentes que têm objetivos comuns no contexto de um ambiente compartilhado (MIETTINEN, 1997). Leont'ev estabelece uma distinção entre ações e atividades:

Do ponto de vista da sua relação com a motivação, um processo interno ou externo aparece ante nós como atividade humana; mas quando o processo é subordinado a um fim, ele aparece como uma ação ou como acumulação de uma cadeia de ações (LEONT'EV, 1978, p. 64).

Desde a perspectiva da teoria da atividade (LEONT'EV, 1978, p. 68), uma atividade interna tem função cognitiva, mas é realizada através de ações externas ou operações psicomotoras. Do mesmo modo, as ações e operações externas podem constituir-se inicialmente como processos internos – cognitivo-fisiológicos – mas sempre mantém a sua integridade como ação ou operação. Portanto, uma separação entre ações cognitivas e motoras não se justifica. Poderíamos pensar as ações psicomotoras como parte de um processo que envolve tanto a retroalimentação epistêmica (processo interno) quanto a atividade pragmática (processo externo).

As ações que não têm por objetivo o resultado sonoro podem servir para obter novos conhecimentos sobre o estado do sistema musical. Essas ações não são independentes dos processos que modificam o ambiente – com resultados imediatos – mas formam um contínuo entre os processos cognitivos e proprioceptivos. Com o apoio da teoria da cognição distribuída (HUTCHINS, 2005), propomos para essa classe de ações a classificação de atividades epistêmicas (KELLER *et al.*, 2009; PIMENTA *et al.*, 2009). Mantendo em mente que os processos epistêmicos e pragmáticos constituem uma unidade sistêmica, podemos diferenciar as ações que têm por objetivo modificar o ambiente – pragmáticas – das que servem para modificar a percepção do nicho ecológico por parte do agente – epistêmicas.

Um tipo importante de atividade são as ações exercidas por um agente sobre os objetos com o objetivo de obter resultados sonoros diretos, ou seja, as ações pragmáticas (KIRSH e MAGLIO, 1994; PIMENTA et al., 2009). As ações pragmáticas têm a função de mudar o ambiente. Já as ações epistêmicas têm um efeito indireto: elas simplificam a realização de ações pragmáticas ao desvendar informações que colocam o agente mais perto da meta a ser atingida. As ações pragmáticas implicam intencionalidade e causalidade. Do ponto de vista da modelagem, elas se encaixam dentro de um fluxo unidirecional de informações, onde cada ação produz um resultado sonoro (ver Figura 1). A maioria dos algoritmos de síntese sonora pertencem a essa classe de modelos (MISRA e COOK, 2009).

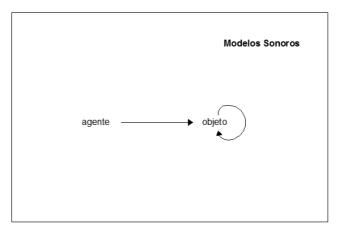


Fig. 1: Modelos sonoros.

## Canais de interação

A constante interação com os objetos e os seres que nos rodeiam molda a nossa percepção do ambiente. O ciclo permanente "ação / percepção / ajuste / nova ação" é o cerne do processo de adaptação do organismo a um novo ambiente. Ou, mais precisamente, a adaptação mútua entre o ambiente e o indivíduo pode ser observada através do conjunto de affordances que surgem a partir deste processo (GIBSON, 1979, p. 127). Ou seja, affordance – ou canal de interação – pode ser definida como sendo qualquer propriedade do ambiente determinada pelas ações do agente ou como o conjunto de propriedades do sistema de interações entre agente e ambiente (CHEMERO e TURVEY, 2007).

A propriocepção acontece no momento em que o organismo interage com o seu contexto através de ações que modificam o ambiente.¹ A atividade proprioceptiva, ou seja, os comandos musculares utilizados para estabelecer contato com

<sup>&</sup>lt;sup>1</sup> Propriocepção é o termo utilizado pela neurociência para descrever os processos perceptivos produzidos através do movimento (KANDEL, 1981; KELSO, 1995).

os objetos, e a influência da percepção desses objetos sobre a atividade em si – através da via aferente – informam o organismo sobre as mudanças que acontecem no ambiente. A atividade proprioceptiva constitui a base para a estruturação das *affordances* (BAERENTSEN e TRETTVIK, 2002). Por esse motivo, *affordance* e atividade estão intrinsecamente interligadas.

A exploração dos canais de interação constitui uma oportunidade de trazer para o campo musical a experiência adquirida nas múltiplas atividades cotidianas desenvolvidas pelo indivíduo ao longo da sua vida. Um paradigma que tem mostrado bons resultados na aplicação de affordances no campo musical é a modelagem ecológica (KELLER, 1999a). Os modelos ecológicos constituem a aplicação em síntese sonora dos processos de interação entre agentes e objetos no meio ambiente. Cada interação gera um evento sonoro, localizado no espaço-tempo. Esse intercâmbio entre agentes e objetos muda o estado do organismo e do material, adaptando características dos objetos às necessidades dos agentes e obrigando-os a mudar o seu comportamento em função do perfil do nicho ecológico. O som é somente um dos produtos dessa interação, portanto a modelagem ecológica é também multimodal (KELLER, 2004) (ver Figura 2).

As atividades musicais modificam o estado das fontes sonoras e dos agentes engajados no fazer musical. Através de um processo seletivo, as ferramentas musicais são adaptadas às necessidades específicas do agente. Simultaneamente, a ação do agente modifica a maneira como ele percebe os objetos utilizados. Em termos mais exatos, as interações entre agente e objeto definem os canais naturais de interação ou *affordances* naturais. Esse conjunto de *affordances* somado aos fatores sociais configura o ambiente pessoal musical.

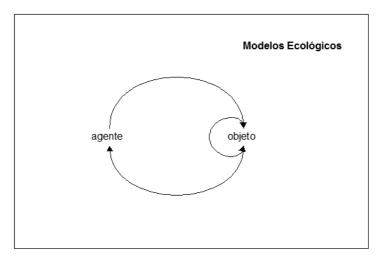


Fig. 2: Modelos ecológicos.

# Níveis de organização sonora

A prática composicional exige a exploração de um amplo leque de resultados sonoros com alto índice de imprevisibilidade. Compor envolve tarefas como categorização, organização, planejamento, seleção, etc. Dependendo do enfoque composicional adotado, as ferramentas de trabalho proporcionam acesso a diferentes níveis de organização do material. Como veremos na discussão sobre as técnicas de síntese sonora, a escolha das ferramentas estabelece um conjunto de variáveis de controle e – juntamente com os materiais – delimita o espaço composicional.

Segundo Keller *et al.* (2009), as competências exigidas pela atividade composicional podem ser definidas nos seguintes termos:

- Imaginário do micro-espaço-tempo: envolve a predição de processos resultantes de variáveis estruturais e transformacionais. As variáveis estruturais descrevem as qualidades sonoras de eventos dentro de uma configuração espacial estática. As variáveis transformacionais informam

sobre as qualidades sonoras dos eventos dentro de configurações espaciais dinâmicas. O imaginário do micro-espaço-tempo engloba as características materiais das fontes sonoras, tais como objetos ressonantes, perfil espectral determinado por algoritmos de síntese, etc. Neste nível é necessária, por exemplo, a projeção dos resultados do processamento digital de sinais e de outros tipos de manipulação do material sonoro.

- Imaginário do meso-espaço-tempo: determina o resultado dos processos no nível mesotemporal, ou seja, levando em conta a interação entre agentes e objetos. Neste nível organizacional entram em jogo variáveis como fase, densidade e distribuição espacial de eventos numa escala temporal superior a aproximadamente 30 milissegundos (KELLER, 1999a).
- Imaginário do macro-espaço-tempo: envolve a predição das relações entre eventos nos vários níveis espaço-temporais, incluindo as interações com o sistema de atividades do grupo social engajado na atividade musical. O imaginário do macro-espaço-tempo que inclui as relações sônico-perceptivas entre eventos temporais é condicionado pelas limitações do sistema auditivo humano. Entre os condicionantes podemos mencionar a memória de curto prazo, a memória de longo prazo e os processos de atenção seletiva (Cf. FORNARI, 2010).

#### Discussão

Concatenando os conceitos apresentados, podemos definir a atividade composicional como ações pragmático-epistêmicas com o objetivo de facilitar projeções micro, meso e macro estruturais (ver Figura 3). Vale destacar que a atividade composicional envolve a constante interação entre agentes e objetos, portanto cada ação modifica tanto o ambiente quanto o sujeito. Dado que o ambiente pessoal resulta da história das interações entre o indivíduo e os nichos ecológicos onde se realizam as atividades cotidianas, e que o processo de adaptação mútua é parcialmente moldado pelos canais sociais de interação, o compositor não pode prever se o seu imaginário vai ser compatível com o imaginário do ouvinte. Duas estratégias podem ser utilizadas como forma de preencher a lacuna entre compositor e ouvinte: 1) a aplicação de nichos

ecológicos artificiais (que funcionam a partir de regras próprias sem seguir padrões encontrados na natureza); 2) a exploração de *affordances* naturais – ou canais naturais de interação - que são comuns à maioria dos nichos ecológicos humanos. Como veremos na seção seguinte, a composição acusmática trabalha com a primeira perspectiva enquanto a ecocomposição geralmente adota a segunda.

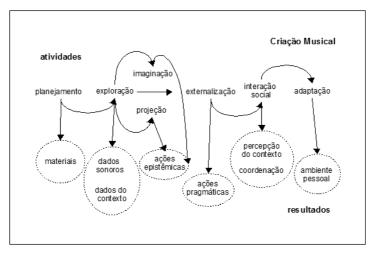


Fig. 3: Composição como atividade.

II

# **Enfoques Composicionais**

partir da conceituação da música como som organizado (VARÈSE, 1959; WEN-CHUNG, 1966), alguns enfoques composicionais têm mudado da manipulação de representações abstratas – com ênfase na notação musical – para a manipulação direta dos parâmetros sonoros. Nesse contexto, as ferramentas computacionais constituem a ponte entre o conceito e o método composicional, visando a obtenção dos resultados musicais. Ao longo dos últimos vinte anos, o aumento na capacidade de processamento dos computadores e a criação de ambientes de programação voltados para a implementação de sistemas em tempo real colocaram ferramentas poderosas de síntese e processamento sonoro ao alcance dos músicos.

Três enfoques que vêm sendo aplicados na criação musical com modelos sonoros, embora claramente separados na sua base conceitual, compartilham alguns métodos de trabalho: as abordagens da música acusmática, da paisagem sonora e da ecocomposição.<sup>2</sup> A música acusmática, fortemente derivada dos princípios da música concreta (musique concrète) expressos por Schaeffer (1966), enfatiza a escuta das características puramente sonoras dos sons como forma de abstrair qualidades musicais para a composição (ver CHION, 2009, p.37). Essa abordagem tem o processamento de sons gravados como uma das principais formas de geração de material musical. A paisagem sonora parte do som ambiental

-

<sup>&</sup>lt;sup>2</sup> Embora existam outras abordagens, a presente discussão se restringirá a essas três. Neste trabalho, optou-se pelo termo "música acusmática" para designar as obras musicais compostas em estúdio, fixadas em suporte e apresentadas em concerto com um conjunto de alto-falantes, que utilizam o objeto sonoro e a escuta reduzida como fundamento teórico-metodológico. No entanto, sabe-se da diversidade de termos possíveis para designar as obras influenciadas esteticamente pelo pensamento de Schaeffer (CHION, 1982).

como matéria prima e eixo estruturador do trabalho composicional, utilizando o processamento sonoro como mecanismo de ampliação da palheta sonora disponível, porém evitando ocultar ou mascarar as fontes sonoras originais. A ecocomposição, por sua vez, focaliza a interação entre agentes e objetos no meio ambiente, utilizando modelos de síntese fundamentados nesse processo como forma de organização do material sonoro. A seguir, discutimos aspectos conceituais desses três enfoques com o intuito de estabelecer similaridades e contrastes nos métodos de manipulação sonora.

### Música acusmática

Um dos aspectos que une as mais diversas produções de música acusmática é que este tipo de música é composta em estúdio e fixada em suporte (fita magnética, CD, DVD-Áudio ou disco rígido), e pensada para ser apresentada numa situação de concerto com um conjunto de alto-falantes – o que geralmente envolve um trabalho de difusão (ou espacialização) dos sons no ambiente .

Ao invés de trabalhar com o conceito de nota musical entendido como representação abstrata dos sons - o compositor de música acusmática interage diretamente com sons gravados, inicialmente escutando-os atentamente e, em seguida, transformando-os e combinando-os em estruturas maiores para constituir uma obra (HARRISON, 1999, p. 118). Assim, em substituição ao trabalho com as relações abstratas das notas musicais, o compositor lida concretamente com os próprios sons. De forma geral, a música acusmática pode ser definida por uma postura estética que se apóia na escuta do compositor, como o juiz supremo das decisões composicionais. Um conceito chave é o de objeto sonoro, entendido como uma entidade sonora coerente e percebida como um todo, abordada de forma isolada do seu contexto original. O objeto sonoro é o correlato da escuta reduzida, uma escuta intencional que se volta à percepção das qualidades puramente sonoras, sem referência à fonte que causou o som ou a qualquer sentido ou conotação exterior às suas características intrínsecas (ver CHION, 2009, p.32). No escopo do trabalho de Schaeffer (1966).

os objetos sonoros são classificados através de um método analítico-descritivo chamado de tipo-morfologia, constituindo, assim, um solfejo (SCHAEFFER e REIBEL, 1967). O compositor de música acusmática extrai, então, estruturas implícitas das configurações explícitas dos objetos (HARRISON, 1999, p. 118). Ou seja, o fazer musical, para Schaeffer, consiste em recolher o concreto sonoro e abstrair as qualidades musicais que o material contém em potência (SCHAEFFER, 1966, p. 23).

O pensamento de Schaeffer, por vezes criticado em certos aspectos (ver WINDSOR, 1995; HOLMES, 2009), deu origem, a uma vasta produção de obras. Embora muitos dos ingredientes do ideário schaefferiano continuem tendo validade e importância para a comunidade acusmática, com o passar dos anos propostas distintas das suas têm florescido mesmo no âmbito dessa comunidade, resultando em estratégias de trabalho mais abertas no que diz respeito à escolha e ao tratamento dos materiais musicais.

Nesse contexto são significativas as considerações de Smalley (1996; 1997), que encara os sons tanto do ponto de vista de suas qualidades puramente sonoras (intrínsecas) quanto de suas possíveis referências e conotações extramusicais (extrínsecas). Os aspectos instrínsecos são analisados e descritos por Smalley (1997) através da espectromorfologia. que, em termos bastante sintéticos, pode ser entendida como uma abordagem voltada à descrição das caraterísticas espectrais dos sons e ao comportamento dinâmico dos mesmos no tempo. Já os aspectos extrínsecos, embora derivem de espectromorfológicas, características estão atrelados questões mais gerais de ordem cultural. Smalley (1997) articula o conceito de ligação com a fonte (source bonding), segundo o qual existiria uma tendência natural a associar os sons a fontes prováveis ou imaginadas e a relacionar sons entre si com base em origens compartilhadas ou aparentes (p. 110), incluindo referências a outras modalidades sensoriais. Os aspectos extrínsecos estão relacionados a uma escuta que Smalley (1996) chama de indicativa, a qual não se restringe a abordar os sons como meros portadores de mensagens ou eventos, mas como parte de um leque de relações que conectam as experiências musicais com as experiências de vida (Smalley, 1996, p. 83).

Pode-se verificar nos estudos recentes sobre música acusmática um crescente interesse por atitudes de escuta não necessariamente associadas à proposta de Schaeffer, mas focadas também no potencial expressivo dos aspectos extrínsecos (ou extra-musicais) associados aos sons (ver SMALLEY, 1996; MANDELBROJT, 1996; GARCIA, 1998; ADKINS, 1999; YOUNG, 2007; KIM, 2008; e BARREIRO, 2010). A obra musical passa, assim, a ser vista não apenas como um sistema auto-referente. Este tipo de enfoque aproxima-se das propostas dos compositores da linha *soundscape*, como veremos na próxima seção.

## Paisagem sonora

Historicamente, as primeiras críticas a uma postura composicional de interferência radical nas características do material musical partiram de compositores da linha *soundscape* ou paisagem sonora (SCHAFER, 2001).<sup>3</sup> A metodologia de separar o material do seu contexto original pode ser descrita como a "objetificação" do som e tem implicações na delimitação do espaço composicional (o contexto musical percorrido na realização de uma obra) e na função da obra dentro do contexto social no qual ela está inserida.

A paisagem sonora inverteu a relação de poder entre materiais e técnicas trazendo o contexto extra-musical para o foco do trabalho criativo. Na abordagem tradicional da paisagem sonora, o ambiente é tanto o ponto de partida quanto o ponto de chegada do processo composicional. Um aspecto importante nesta linha de trabalho é a fidelidade às fontes e a contextualização cuidadosa através de elementos extramusicais. Poderíamos dizer que o espaço geográfico é tanto o marco estrutural quanto a fonte exclusiva do material da obra.

Neste enfoque, a intervenção do compositor é reduzida ao mínimo e – nas formas mais puras – as relações entre as

<sup>-</sup>

<sup>&</sup>lt;sup>3</sup> A palavra *soundscape* é um neologismo cunhado por R. M. Schafer a partir da junção de *sound* (som) com *landscape* (paisagem). A tradução literal é paisagem sonora.

fontes sonoras originais permanecem quase intocadas (WESTERKAMP, 1998). O compositor utiliza material gravado e edita as gravações tentando manter as características temporais e espaciais das fontes sonoras encontradas no local. Assim, o ouvinte é convidado a recriar o contexto do qual os materiais foram extraídos.

Embora haja uma tendência a restringir a interferência do compositor, é inegável que há escolhas composicionais nessa abordagem, as quais são feitas durante todo o processo de coleta, manipulação e organização sonora. As escolhas dos locais de gravação, do equipamento, do posicionamento dos microfones, dos métodos de edição, dos formatos de apresentação, dos suportes de mídia e da relação com o público-alvo são decisões composicionais que têm consequências estéticas e políticas (ver discussão em KELLER e CAPASSO, 2001). Neste ponto é que a ecocomposição toma distância da paisagem sonora.

# Ecocomposição

Com o intuito de abrir caminhos para a interação público-obra, a ecocomposição propõe procedimentos de trabalho que fomentam a organicidade do material e do contexto da obra. A metodologia ecocomposicional inclui a acumulação e a modularidade como processos estruturadores do material sonoro, a interação como mecanismo para adaptar o material às ações do público no espaço da obra, a vivência da temática da obra pelos artistas-realizadores, e a ancoragem, isto é, a relação dialógica entre o ambiente e o processo composicional adotado (KELLER, 2000; KELLER e CAPASSO, 2006).

Materiais, métodos, formato e audiência, todos são parte integrante do processo ecocompositional. Enquanto a música acusmática tem o compositor como figura-chave e a paisagem sonora tem o som ambiental como o elemento mais importante do trabalho, a ecocomposição coloca a ênfase na experiência composicional em si. O conceito chave é o da vivência (o termo utilizado na literatura em inglês é reenaction). Em primeiro lugar, o ecocompositor define um

conceito historicamente e geograficamente localizado. Então, esse conceito é colocado em prática, em outras palavras, a experiência histórico-social é revivida no seu ambiente de origem. Enquanto é realizada a experiência ecocomposicional, o material e o contexto colocam novas relações em jogo obrigando a reformular ou possivelmente mudar os axiomas iniciais da proposta. Na verdade, um termo mais exato do que "axiomas" é "condições iniciais" do sistema ecocomposicional. Desse modo, o método de composição torna-se um sistema aberto que é modificado pelas interações entre os agentes e os objetos dentro de um espaço geográfico e temporal prédefinido. Essa abertura estrutural permeia todo o processo de composição, desde a coleta das fontes sonoras até o formato de suporte da obra. Portanto as ecocomposições geralmente são obras modulares e abertas que dependem do público para atingir a sua forma final. Levada às últimas consequências, uma ecocomposição não só deve ser moldada pelo processo criativo, também deve mudar a maneira em que o compositor percebe o foco conceitual de sua obra. E. em última instância, deve criar pontes entre o público-participante e a realidade mediada pelo trabalho artístico para propiciar a interação.

#### Discussão

Retomando a questão das interações entre ações pragmáticas e epistêmicas, mencionada na primeira parte deste trabalho, podemos afirmar que, durante a composição de uma obra, o compositor gradativamente delimita o seu espaço composicional através da coleta de material, da escolha das variáveis de controle nos processos de manipulação sonora e da sua vivência antes e durante a atividade composicional. Os percursos por esse espaço condicionam não apenas a criação e modelagem dos materiais, mas também a seleção e combinação dos mesmos em unidades maiores que, por fim, resultam na obra musical. Esse resultado pode ser tanto uma obra em suporte fixo quanto um campo de possibilidades sonoras a serem articuladas através da interação com outros indivíduos, sejam eles participantes de uma instalação ou músicos instrumentistas.

Em termos amplos, podemos dizer que a composição acusmática consiste na manipulação de objetos sonoros através de operações que visam abstrair qualidades musicais do material gravado. A ênfase é colocada no papel do compositor, permitindo estabelecer relações que podem ser sugeridas pelo próprio material ou, em abordagens mais abertas, que podem ser determinadas pelo mapeamento de parâmetros extrasonoros. O resultado musical desse processo é a obra fixada em suporte a ser articulada no ambiente de concerto através de alto-falantes.

Já a paisagem sonora coloca o ambiente sonoro no centro do trabalho composicional, enfatizando o contexto ao qual o material sonoro pertence, e colocando as técnicas de manipulação sonora em função da reinserção dos elementos extra-sonoros encontrados no ambiente. Paradoxalmente, esse enfoque nem sempre demanda a interação do músico com o ambiente local. Portanto, é possível realizar uma obra a partir do acesso às fontes sonoras utilizando arquivos como o existente no *World Soundscape Project*, da Simon Fraser University (TRUAX, 2002). Em contraste com os métodos ecocomposicionais, as composições de paisagens sonoras geralmente não exigem a interação com o público para atingir a sua forma final. O trabalho é concluído no momento em que o compositor finaliza a edição do material sonoro.

Em linha com o enfoque de paisagem sonora, a ecocomposição mantém a relação entre o material e o contexto, mas aprofunda a ação composicional ao estabelecer mecanismos de manipulação do material sonoro baseados nos eventos observados na interação agente-objeto. Além disso, em contraste com a abordagem acusmática, o espaço onde ocorre a interação entre esses diversos elementos transborda o ambiente do estúdio de composição, incorporando o contexto geográfico com o qual a obra dialoga.

#### Ш

# Exemplos de implementações e realizações musicais

Nesta seção focalizamos o nosso trabalho composicional por duas razões: 1) como compositores, temos conhecimento de primeira mão dos detalhes técnicos e do contexto no qual as obras foram realizadas, 2) existem publicações em português sobre o trabalho de compositores como R. M. Schafer e Pierre Schaeffer, mas a discussão da obra de compositores mais recentes é escassa. Os conceitos abordados relativos à síntese granular estão definidos no glossário deste livro. Sugerimos a leitura desses conceitos antes do estudo dos exemplos composicionais.

# 112 El Escrache / sCRATch

toco y me voy / touch'n'go (KELLER, 1999b) é uma obra multimídia para ator bilíngue e fita octofônica (ver a disposição do sistema de alto-falantes na Figura 4). O material sonoro inclui sons gravados e sons sintetizados a partir de modelos ecológicos (KELLER, 1999a). toco y me voy propõe um sistema organizacional que demanda um formato modular e flexível. Cada módulo da obra é uma peça, sendo que os módulos podem ser combinados em quatro seções de dez a dezoito minutos de duração. O eixo estruturador é refletido no tema da obra – o "gaúcho" ou viageiro –, na sua estrutura – um caminho definido pelo participante-ouvinte –, e pelo material sonoro utilizado – com referências diretas ao contexto social. As ideias desenvolvidas no texto da obra – e apresentadas em formato multimídia – são estruturadas em módulos correspondentes a cada seção sonora.

O material da obra foi produzido a partir de técnicas de síntese resultantes da interação entre processos de ativação e corpos ressonantes. A unidade sônica é o evento. Quer dizer não há sons sem ataque ou sem queda natural. No mesotempo, todos os algoritmos utilizados seguem padrões ecológicos. Ou

seja, eles produzem eventos reais ou possíveis no âmbito terrestre. No macrotempo, a peça explora as propriedades emergentes da interação de processos nos níveis micro e médio. A forma final da obra depende da exploração intencional do ouvinte através dos caminhos sugeridos no texto e nos módulos sonoros.

Um dos módulos da obra exemplifica de forma clara a relação entre técnicas de síntese e conceitos extra-sonoros. *El Escrache / sCRATch* estabelece um jogo de significados entre a ação de friccionar (*scratch* em inglês) e a gíria portenha "escrachar". O movimento de direitos humanos da Argentina adotou esse termo para se referir à ação de colocar em evidência a presença em locais públicos de torturadores e assassinos que não foram condenados pela justiça. O "escrache" consiste numa ação pública para chamar a atenção ao fato de que criminosos compartilham direitos e atribuições dos cidadãos que não cometeram atos criminais.

O material sonoro de El Escrache / sCRATch foi gerado a partir de amostras sonoras e sons sintetizados aplicando um modelo de fricção (ver KELLER, 1999a). A fricção é geralmente vinculada aos gestos produzidos por agentes humanos. Portanto, os eventos de fricção são limitados a durações que podem ser efetuadas pelo movimento do braco humano. A ação de raspagem não pode iniciar com níveis altos de energia, mas precisa se desenvolver lentamente partindo da amplitude zero. aumentando até o máximo de energia, até voltar ao ponto de repouso. Para imitar esse comportamento, utilizamos máscaras de tendência que permitem variações aleatórias na amplitude dos grãos mas que limitam os estratos iniciais e finais à amplitude zero. Da mesma forma, o aumento e a diminuição da velocidade na ação de raspar acontece de forma relativamente lenta. A simulação do processo de ativação foi feita controlando a taxa de iteração de grãos através de um algoritmo de acumulação.

O módulo de ativação de fricção consiste em um gerador de números aleatórios que envia valores distribuídos uniformemente em torno do zero e de um acumulador com limites em zero e um. Esses limites estabelecem as velocidades de fricção mais lentas (aproximando-se a zero) e as velocidades

mais rápidas (nos valores perto de um). A velocidade aumenta ou diminui de acordo com o número de valores positivos ou negativos produzidos pelo gerador de números aleatórios. O valor-limite absoluto dado ao gerador de números define uma forma de quantização temporal. Se esse valor é alto aumenta o intervalo temporal entre os grãos, se o valor é baixo, a taxa de iteração granular aumenta simulando ações de fricção rápida.

Além da ação do agente, as características sonoras da raspagem também são determinadas pelo tipo de superfície a ser friccionada. Dado que o gesto friccional é considerado como relativamente constante, as irregularidades na taxa de iteração entre o objeto ativador e a superfície podem ser atribuídas à rugosidade da superfície. Através da randomização da taxa de iteração de grãos dentro de uma pequena porcentagem do valor gerado pelo módulo de ativação (até 10%), obtém-se diferentes graus de rugosidade sem afetar a causalidade gestual do modelo. O material da superfície a ser raspada é estabelecido pela qualidade timbrística da coleção de grãos ou manipulando o resultado espectral através da utilização de ressonadores.

# Maresia, Sons Adentro e I/VOID/O

*Maresia* (BARREIRO, 2006b) e *Sons Adentro* (BARREIRO, 2006c) são obras eletroacústicas em oito canais (fita octofônica), utilizando a disposição de alto-falantes ilustrada na Figura 4.

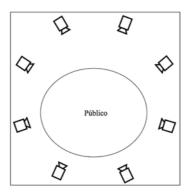


Fig. 4: Posicionamento do público e o sistema de oito canais.

Ambas as obras utilizam sons produzidos com técnicas granulares através de um aplicativo em Max/MSP adaptado por Barreiro (2006a) a partir de uma versão para doze canais desenvolvida por Peter Batchelor e baseada em um granulador estéreo de Erik Oña.4 O aplicativo de Batchelor inclui a possibilidade de percorrer a forma de onda do som prégravado com o mouse durante o processo de granulação definindo, em tempo real, as porções do arquivo a serem utilizadas para a amostragem dos grãos. Além disso, apresenta um dispositivo que permite a configuração dos canais a serem utilizados na distribuições espaciais aleatórias. A versão de Barreiro (2006a), para oito canais, apresenta alguns recursos adicionais: a) um dispositivo automático que percorre o arquivo pré-gravado em loop com uma velocidade controlada pelo usuário ou determinada aleatoriamente - definindo, a cada passo, o ponto de amostragem dos grãos<sup>5</sup>; b) um seguidor de envelope (envelope follower) que muda a velocidade de varredura do arquivo pré-gravado quando o som gerado

<sup>&</sup>lt;sup>4</sup> Batchelor realizou modificações subsequentes em seu aplicativo – atualmente chamado de Plode – cuja versão mais recente é para oito canais (ver http://www.peterbatchelor.com/maxPlode.html).

<sup>&</sup>lt;sup>5</sup> Conforme dito anteriormente, na implementação de Batchelor este parâmetro dependia da manipulação manual do usuário.

apresenta picos de amplitude acima de um limiar préestabelecido. $^6$ 

A abordagem composicional utilizada em *Maresia* e *Sons Adentro* segue, em grande parte, os princípios da música acusmática mencionados na segunda parte deste trabalho. Em certos momentos, entretanto, há a adoção de um leque mais amplo de materiais sonoros – incluindo sons ambientais de origem reconhecível – o que revela traços de uma abordagem do tipo paisagem sonora entremeada com a abordagem acusmática preponderante.

Em *Maresia* – baseada principalmente em sons do mar – uma seção inteira (de 3'07" a 4'12") utiliza sons de bolhas de ar na água, que são granulados e distribuídos nos oito canais.<sup>7</sup> Esse trecho é seguido por uma seção contrastante (de 4'13" a 6'53") que também faz uso de sons gerados pela granulação de sons de cantos de pássaros. Esses sons tiveram suas durações estendidas (ver *time stretching*), revelando ocasionalmente uma morfologia interna de caráter granular iterativo.

Sons Adentro – baseada num variado leque de fontes sonoras – apresenta vários sons de morfologia iterativa, tais como bolinhas pingando, sons de grilos e sons gerados por

<sup>&</sup>lt;sup>6</sup> Lippe (1994) menciona a possibilidade de derivar o controle dos parâmetros de granulação da análise de características do sinal sonoro de instrumentos em obras para meios acústicos e eletrônica em tempo real. A abordagem adotada aqui é ligeiramente distinta, uma vez que o próprio som gerado pelo aplicativo é que é analisado, instituindo,

assim, uma forma de retroalimentação dentro do processo de síntese. A definição automática de parâmetros através de procedimentos desse tipo tem o interessante aspecto de diminuir a dificuldade do controle simultâneo de múltiplas variáveis de granulação (ver LIPPE, 1994; MILETTO et al., 2004).

<sup>&</sup>lt;sup>7</sup> Ver o módulo *Palabras...* da obra *toco y me voy* (KELLER, 1999b) para um exemplo de uso do mesmo material sonoro, gerado a partir de modelos ecológicos.

colisões repetidas entre ímãs de formato amendoado.<sup>8</sup> A morfologia iterativa foi utilizada como uma maneira de estabelecer relações entre sons de origem reconhecível e sons de caráter mais abstrato. Alguns desses sons são apresentados sem processamento. Outros têm seu caráter iterativo potencializado pela granulação.

I/VOID/O (ABREU e BARREIRO, 2008; BARREIRO et al., 2009) é uma instalação que apresenta imagens e sons gerados em tempo real como resultado de interações com os visitantes. Os visitantes se deparam com uma esfera metálica cujo interior espelhado pode ser observado através de imagens captadas por câmeras e exibidas em dois telões. Uma das câmeras localizadas no interior da esfera encontra-se posicionada na extremidade de uma haste cujo segmento externo pode ser manipulado pelos visitantes. A interação com a esfera, através da manipulação da haste, propicia a passagem por diferentes formas de observação do seu interior (níveis de imersão), gerando sons que acompanham - por analogia comportamento das imagens. Em alguns momentos, os sons são gerados pela granulação de amostras de áudio pré-gravado através de dois granuladores independentes que funcionam em paralelo produzindo até 20 estratos granulares a partir de três arquivos de áudio diferentes. Os grãos sonoros são distribuídos aleatoriamente num sistema de quatro canais, constituindo um ambiente sonoro imersivo. Através de configurações específicas do aplicativo de granulação, foram buscados resultados sonoros que não apresentassem semelhanças perceptíveis com as fontes sonoras originais, aproximando-se de uma abordagem acusmática mais tradicional. No entanto, a geração de material sonoro em tempo real através de um processo interativo instaura uma postura composicional distinta da abordagem acusmática - em que o material sonoro é trabalhado em estúdio e fixado em suporte para a apresentação da obra em concerto e mais próxima da abordagem ecocomposicional.

<sup>&</sup>lt;sup>8</sup> Para uma comparação do uso de colisões a partir de modelos ecológicos, ver o módulo *Pandemonium* da obra *toco y me voy* (KELLER, 1999b).

Nos três trabalhos mencionados acima, a opção pelo uso de técnicas granulares foi motivada pela riqueza timbrística dos sons gerados e pela possibilidade de explorar diferentes tipos de morfologia sonora dependendo da configuração dos parâmetros de granulação. No que diz respeito à distribuição espacial dos sons, os resultados obtidos variam entre sons que se apresentam de forma difusa e outros que são localizados em pontos específicos do espaço. Isso depende das características espectrais dos sons granulados, dos valores utilizados nos parâmetros de granulação e do controle do número de canais envolvidos na distribuição espacial dos grãos. Embora o posicionamento dos grãos seja feito de forma aleatória, o controle do número de canais envolvidos pode resultar em sons que descrevem trajetórias com um caráter direcional. Isso pode ser verificado, por exemplo, no som inicial de Sons Adentro, que descreve um percurso em ziguezague (errático), mas com uma direcionalidade que parte dos alto-falantes frontais em direção aos alto-falantes traseiros.

# 118

#### Discussão final e desenvolvimentos futuros

As atividades composicionais incluem, mas não se limitam, a planejamento, exploração, imaginação, projeção, externalização, interação social e adaptação. Através de um processo de adaptação mútua, essas atividades envolvem interações entre os agentes e os objetos e são moldadas por canais naturais e sociais de interação. O processo criativo fomenta o desenvolvimento de ferramentas, o estabelecimento de vínculos sociais e a adoção de novas formas de organização sonora.

Na discussão prévia, observamos que as técnicas granulares aplicadas no contexto dos enfoques da música acusmática, da paisagem sonora e da ecocomposição fornecem acesso a parâmetros composicionais fundamentais para a manipulação do material sonoro. No entanto, o desenvolvimento técnico ficaria vazio se não houvesse um suporte conceitual para guiar o trabalho de implementação. Paradoxalmente, esse avanço técnico muitas vezes impulsiona a convergência de conhecimentos dispersos em diferentes áreas

gerando suporte experimental para novos conceitos. Esse é caso da síntese granular tátil – dentro da abordagem ecológica – e da síntese com base em modelos bio-inspirados – na composição acusmática e em sistemas interativos em tempo real. Os dois novos eixos de trabalho explorados nessas vertentes são a multimodalidade (particularmente o tato) na síntese granular tátil e a correlação entre parâmetros sonoros e o comportamento coletivo auto-organizado, no caso da síntese com modelos bio-inspirados.

A síntese granular tátil adiciona a sensação proprioceptiva ao processo de interação com dados sonoros através de transdutores táteis (BARRASS e ADCOCK, 2002). A sensação tátil correlacionada à síntese de eventos sonoros fornece a oportunidade de estabelecer nexos entre as duas modalidades perceptivas. Desta feita, a exploração sonora é diretamente espelhada na exploração de texturas concretas, ampliando o espaço composicional ao campo dos objetos bi e tridimensionais. Compor se transforma literalmente em esculpir o som.

A síntese sonora baseada em modelos bio-inspirados fundamenta-se no fato de que, na natureza, estruturas complexas surgem como fenômeno global emergente gerado a partir de interações simples e localizadas realizadas entre indivíduos de uma colônia de animais. Na área de inteligência artificial, Carvalho et al. (2004) mencionam os estudos em inteligência de enxames (swarm intelligence) destinados à solução de problemas complexos - particularmente problemas de otimização combinatória - através da implementação de algoritmos inspirados em colônias de formigas - Ant Colony Optimization (ACO) e Ant Clustering Algorithm (ACA). Embora as capacidades de uma formiga isolada sejam limitadas, as ações coletivas da colônia propiciam a realização de tarefas sofisticadas - como é o caso da identificação do menor caminho entre o formigueiro e uma fonte de alimentos (CARVALHO et al., 2004, p. 22). Isso ocorre com base no processo de estigmergia que propicia a comunicação indireta que os insetos sociais manifestam ao alterar o ambiente com o depósito de (GRASSÉ, 1959). estigmergia pode feromônios Α diferenciada de outras formas de comunicação através de dois fatores: 1) o caráter físico da informação corresponde à

modificação dos estados físicos do meio-ambiente visitado pelos insetos e 2) o acesso local à informação, somente disponível para os insetos que visitam o lugar onde a informação foi criada. O mesmo conceito pode ser aplicado ao âmbito computacional, quando partes do sistema se comunicam de forma indireta modificando o seu ambiente local (GRASSÉ, 1959).9

O algoritmo Boids, de Craig Reynolds (2009), é outro exemplo de algoritmo inspirado em processos naturais autoorganizados, simulando os movimentos de um conjunto de animais - tais como bandos de pássaros - a partir de regras bastante simples. Reynolds (2009) define o procedimento algorítmico da seguinte forma: a) evitar colisões com indivíduos vizinhos; b) manter aproximadamente a mesma direção e velocidade que os indivíduos vizinhos; c) não se distanciar muito dos outros membros do grupo. Tomados em conjunto, os movimentos apresentam alta coerência, apesar de serem realizados por movimentos individuais bastante simples. Swarm Granulator (BLACKWELL e YOUNG, 2004a) é um sistema computacional baseado no algoritmo Boids que interage com instrumentos musicais para gerar música improvisada através de síntese granular.10 implementação, os boids tendem a se movimentar em torno de atratores posicionados num espaço de seis dimensões com base em parâmetros musicais extraídos da análise de sinais sonoros dos instrumentos musicais. As posições das partículas são reinterpretadas como parâmetros e esses parâmetros são utilizados para modelar novos sons através de síntese granular.

A abordagem bio-inspirada ocorre também na síntese granular com sistemas multicanais. A distribuição dos grãos no espaço pode ser realizada não só de forma aleatória (dentro de limites pré-estabelecidos) mas também pode ser definida através de métodos algorítmicos, como é o caso nas

<sup>&</sup>lt;sup>9</sup> Para um exemplo de um sistema de composição musical inspirado em colônias de formigas, ver GUÉRET, MONMARCHÉ e SLIMANE (2004).

<sup>&</sup>lt;sup>10</sup> Ver também Blackwell e Young (2004b).

implementações de enxames de partículas (KIM-BOYLE, 2005) e na aplicação do algoritmo *Boids* (KIM-BOYLE, 2006; WILSON, 2008). Nesses casos, as trajetórias das partículas são mapeadas para os movimentos dos grãos sonoros no espaço, possibilitando a criação de trajetórias com padrões complexos – apontando, assim, interessantes possibilidades para o controle dos processos de difusão sonora com sons granulares.

Fenômenos emergentes e auto-organizados, similares aos citados acima, ocorrem na improvisação de música experimental quando não existem regras de largo alcance definidas a priori (LEWIS, 2007). O que se obtém com esses processos são sistemas que interagem de forma integrada com os músicos humanos através de processos de controle de parâmetros sonoros. O resultado sonoro global (a macroestrutura) surge como fenômeno emergente derivado da interação local entre os improvisadores. Os músicos e os processos algorítmicos moldam - mas não determinam - os elementos micro-estruturais no âmbito do material sonoro. Essa estratégia de interação entre agentes é a característica principal dos sistemas com propriedades emergentes. Tanto a abordagem ecológica quanto os modelos bio-inspirados compartilham o mesmo espaço composicional aberto e apontam para desenvolvimentos que podem transformar profundamente o que entendemos por criação musical.

# Agradecimentos

O primeiro autor agradece pelo apoio da Capes (Bolsa de Doutorado no Exterior) e do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (Bolsa de Pós-Doutorado Júnior) com as quais foram obtidos alguns dos resultados apresentados neste trabalho. A pesquisa do segundo autor foi financiada pelo CNPq nos projetos 301982/2008-2, 478092/2008-3 e 571443/2008-7.

<sup>&</sup>lt;sup>11</sup> Ver também Davis e Rebelo (2005) para uma abordagem sobre a distribuição de sons no espaço utilizando o algoritmo *Boids*.

#### Referências

ABREU, Sérgio Canavazzi de; BARREIRO, Daniel Luís. *I/VOID/O*. Instalação interativa concebida e programada por Sandro Canavezzi de Abreu, com eventos sonoros gerados em tempo real através de aplicativo programado por Daniel Barreiro, 2008. Registro em vídeo disponível em: http://www.youtube.com/watch?v=fnffoU7pX2o

ADKINS, Monty. Acoustic chains in acousmatic music. In: *Imaginary Space: Proceedings of the 1999 Australasian Computer Music Conference*. Wellington: University of Wellington, 1999. Disponível em: http://eprints.hud.ac.uk/4273

BAERENTSEN, Klaus; TRETTVIK, Johan. An activity theory approach to affordance. In: *Proceedings of the Second Nordic Conference on Human-Computer Interaction* (NORDI-CHI 2002), Aarhus, Dinamarca, 2002, p. 51–60.

BARRASS, Stephen; ADCOCK, Matt. Interactive granular synthesis of haptic contact sounds. In: *Proceedings of the 22nd International Conference on Virtual, Synthetic and Entertainment Audio*, Espoo, Finlândia, 2002.

BARREIRO, Daniel Luís. *Portfolio of Compositions*. Birmingham, 2006a. Tese (Doutorado) University of Birmingham.

BARREIRO, Daniel Luís. *Maresia*. Obra acusmática em oito canais (2005-2006). Versão estéreo no CD *Metamorphoses 2006*. Bruxelas: Musiques&Recherches, MR-2006-CD, 2006b.

BARREIRO, Daniel Luís. *Sons Adentro*. Obra acusmática em oito canais. Mídia digital, 2006c.

BARREIRO, Daniel Luís. Sonic image and acousmatic listening. *Organised Sound*, v. 15, n. 1 (2010), p. 35-42.

BARREIRO, Daniel Luís; ABREU, Sérgio Canavazzi de; CARVALHO, André Ponce de León F. de. I/VOID/O: real-time sound synthesis and video processing in an interactive installation. In: *Proceedings of the SBCM 2009* – 12th Brazilian Symposium on Computer Music, Recife, 2009. Disponível em: http://compmus.ime.usp.br/sbcm/2009/proceedings-SBCM2009.pdf

BLACKWELL, Tim; YOUNG Michael. Swarm Granulator. In: *Applications of Evolutionary Computing: Proceedings of the EvoWorkshops 2004: EvoBIO, EvoCOMNET, EvoHOT, EvoISAP, EvoMUSART, and EvoSTOC,* Coimbra, Portugal, p. 399-408. Berlim / Heidelberg: Springer, 2004a. Série "Lecture Notes in Computer Science", v. 3005. Disponível em: http://www.doc.gold.ac.uk/~mas01tb/papers/SwarmGranulator.pdf

BLACKWELL, Tim; YOUNG Michael. Self-organised music. *Organised Sound*, v. 9, n. 2 (2004b), p. 123-136. Disponível em: http://www.doc.gold.ac.uk/~mas01tb/papers/SelfOrganisedMusic.pdf

CARVALHO, André Carlos P. L. Ferreira de; DELBEM, Alexandre Cláudio Botazzo; SIMOES, Eduardo do Valle; TELLES, Guilherme Pimentel; ROMERO, Roseli Aparecida Francelim. *Computação Bioinspirada*. Apostila de minicurso XXIII JAI - Congresso da SBC. Porto Alegre: Sociedade Brasileira de Computação, 2004.

CHEMERO, Anthony; TURVEY, Michael T. Gibsonian affordances for roboticists". *Adaptive Behavior*, v. 15, n. 4 (2007), p. 473–480.

CHION, Michel. La musique électroacoustique. Paris: PUF, 1982.

\_\_\_\_\_. Guide to sound objects. Pierre Schaeffer and musical research.
Trad. John Dack e Christine North, 2009. Disponível em:
http://www.ears.dmu.ac.uk/spip.php?page=articleEars&id\_article=359

DAVIS, Tom; REBELO, Pedro. Hearing emergence: towards sound-based self-organisation. In: *Proceedings of ICMC 2005* – International Computer Music Conference, Barcelona, 2005, p. 463–466.

FORNARI, José. Percepção, cognição e afeto musical. In: KELLER, Damián; BUDASZ, Rogério (orgs.). *Criação musical e tecnologias: teoria e prática interdisciplinar*, p. 6-38, Goiânia: ANPPOM, 2010.

GARCIA, Denise Hortência Lopes. *Modelos perceptivos na música eletroacústica*. São Paulo, 2008. Tese (Doutorado) Pontifícia Universidade Católica de São Paulo.

GIBSON, James J. *The ecological approach to visual perception*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1979.

GRASSÉ, Pierre-Paul. La reconstruction du nid et les coordinations interindividuelles. La théorie de la stigmergie. *Insectes Sociaux*, v.6, 1959, p. 41-84.

GUÉRET, Christelle; MONMARCHÉ, Nicolas; SLIMANE, Mohamed. Ants can play music. In: *Ant colony optimization and swarm intelligence: Proceedings of the 4th International Workshop ANTS 2004*, Bruxelas, p. 310-317. Berlim / Heidelberg: Springer, 2004.
Série "Lecture Notes in Computer Science", v. 3172. Disponível em:

http://www.hant.li.univ-tours.fr/webhant/pub/GueMonSli04a.ants.pdf

HARRISON, Jonty. Sound, space, sculpture: some thoughts on the "what", "how" and "why" of sound diffusion. *Organised Sound*, v. 3, n. 2 (1999), p.117-127.

HOLMES, Bryan. Análise espectromorfológica da obra eletroacústica "Desembocaduras". *Revista Eletrônica de Musicologia*, v. 12 (2009). Disponível em:

http://www.rem.ufpr.br/\_REM/REMv12/02/analise\_morfologica\_embo caduras.htm

HUTCHINS, Edwin. Material anchors for conceptual blends. *Journal of Pragmatics*, v. 37 (2005), p. 1555-1577.

KANDEL, Eric R. Somatic sensory system III: central representation of touch. In: KANDEL, Eric R.; SCHWARTZ, James H. (orgs.). *Principles of neural science*. Nova York: Elsevier, 1981.

KELLER, Damián. touch'n'go: Ecological Models in Composition. Burnaby, Canadá, 1999a. Dissertação (Mestrado em Arte) Simon Fraser University. Disponível em: http://ccrma.stanford.edu/~dkeller

KELLER, Damián. *touch'n'go / toco y me voy* [CD-ROM]. Burnaby, Canadá: Earsay Productions, 1999b. Disponível em: http://www.earsay.com/soundshop/earsay/CDs/tng.html

KELLER, Damián. Compositional processes from an ecological perspective. *Leonardo Music Journal*, v. 10 (2000), p. 55-60. Disponível em:

http://www.mitpressjournals.org/doi/abs/10.1162/09611210057045

KELLER, Damián. *Paititi: a multimodal journey to El Dorado*. Stanford, CA, 2004. Tese (Doutorado em Artes Musicais) Stanford University. Disponível em: http://portal.acm.org/citation.cfm?id=1048456

KELLER, Damián; BARROS, Ana Elisa Bonifácio; FARIAS, Flávio Miranda de; NASCIMENTO, Rafael Vasconcelos do; PIMENTA, Marcelo Soares; FLORES, Luciano Vargas; MILETTO, Evandro Manara.; RADANOVITSCK, Eduardo Aquiles Affonso; SERAFINI, Rafael de Oliveira; BARRAZA, José F. Música ubíqua: conceito e motivação. In: *Anais do XIX Congresso da ANPPOM*, Curitiba, 2009, p. 539-542.

KELLER, Damián; CAPASSO, Ariadna. Social and perceptual processes in the installation "The Trade". *Organised Sound*, v. 5 (2001), p. 85-94.

KELLER, Damián; CAPASSO, Ariadna. New concepts and techniques in eco-composition. *Organised Sound*, v. 11 (2006), p. 55-62.

KELSO, J. A. Scott. *Dynamic patterns: the self-organization of brain and behavior*. Cambridge, MA: MIT Press, 1995. Disponível em: http://books.google.com/books?id=zpjejjytkiIC&pgis=1

KIM, Suk-Jun. *Listeners and imagination: a quaternary framework for electroacoustic music listening and acousmatic reasoning.* Gainesville, 2008. Tese (Doutorado) University of Florida.

KIM-BOYLE, David. Sound spatialization with particle systems. In: *Proceedings of the 8th International Conference on Digital Audio Effects* (DAFX-05), Madri, 2005, p. 65-68.

KIM-BOYLE, David. Spectral and granular spatialization with boids. In: *Proceedings of ICMC 2006* – International Computer Music Conference, New Orleans, 2006, p. 139-142.

KIRSH, David; MAGLIO, Paul. On distinguishing epistemic from pragmatic action. *Cognitive Science*, v. 18 (1994), p. 513–549.

LAZZARINI, Victor. Técnicas fundamentais para instrumentos computacionais de síntese e processamento de áudio. In: KELLER, Damián; BUDASZ, Rogério (org.). *Criação musical e tecnologias: teoria e prática interdisciplinar*, p. 39-69, Goiânia: ANPPOM, 2010.

LEONT'EV, Aleksei Nikolaevich. *Activity, consciousness, and personality*. Englewood Cliffs: Prentice-Hall, 1978.

LEWIS, George E. Live algorithms and the future of music. CT Watch Quarterly, v. 3, n. 2 (mai. 2007). Disponível em: http://www.ctwatch.org/quarterly/articles/2007/05/live-algorithms-and-the-future-of-music/index.html

LIPPE, Cort. Real-time granular sampling using the IRCAM signal processing workstation. *Contemporary Music Review*, v. 10 (1994), p. 149-155. Disponível em:

http://www.music.buffalo.edu/faculty/lippe/pdfs/granular.pdf

MANDELBROJT, Jacques (org.). Les unites sémiotiques temporelles. Marselha: Laboratoire Musique et Informatique de Marseille, 1996.

MIETTINEN, Reijo. The concept of activity in the analysis of heterogeneous networks in innovation process. In: *CSTT Workshop 'Actor Network and After'*, julho de 1997.

MILETTO, Evandro Manara; COSTALONGA, Leandro Lesqueves; FLORES, Luciano Vargas; FRITSCH, Eloi Fernando; PIMENTA, Marcelo Soares; VICARI, Rosa Maria. *Introdução à Computação Musical*, 2004. Disponível em: http://www.evandromanara.net/files/UFRGS/MinicursoIntrodCM\_CBC2004.pdf

MILETTO, Evandro Manara; PIMENTA, Marcelo Soares; BOUCHET, François; SANSONNET, Jean-Paul; KELLER, Damián. Music creation by novices should be both prototypical and cooperative: lessons learned from CODES. In: *Proceedings of the 12th Brazilian Symposium on Computer Music*, Recife: SBC, 2009. Disponível em: http://gsd.ime.usp.br/sbcm/

MISRA, Ananya; COOK, Perry R. Toward synthesized environments: a survey of analysis and synthesis methods for sound designers and

composers. In: Proceedings of ICMC 2009 – International Computer Music Conference.

PIMENTA, Marcelo Soares; FLORES, Luciano Vargas; CAPASSO, Ariadna; TINAJERO, Patrícia; KELLER, Damián. Ubiquitous music: concept and metaphors. In: *Proceedings of the XII Brazilian Symposium on Computer Music*, Recife: SBC, 2009. Disponível em: http://gsd.ime.usp.br/sbcm/

REYNOLDS, Craig. *Boids* [Página do autor], 2009. Disponível em: http://www.red3d.com/cwr

SCHAEFFER, Pierre. *Traité des objets musicaux*. Paris: Éditions du Seuil, 1966.

SCHAEFFER, Pierre; REIBEL, Guy. Solfège de l'objet sonore. Paris: ORTF, 1967.

SCHAFER, R. Murray. A afinação do mundo. São Paulo: Ed. UNESP, 2001.

SMALLEY, Denis. The listening imagination: listening in the electroacoustic era. *Contemporary Music Review*, v.13, n. 2 (1996), p. 77-107.

SMALLEY, Denis. Spectromorphology: explaining sound-shapes. *Organised Sound*, v. 2, n. 2 (1997), p.107-126.

TRUAX, Barry. Genres and techniques of soundscape composition as developed at Simon Fraser University. *Organised Sound*, v. 7, n. 1 (2002), p. 5-14.

VARÈSE, Edgar. *Rhythm, form and content*, 1959. [Palestra - excertos publicados em *The Liberation of Sound*]. In: CHILDS, Barney; SCHWARTZ, Elliott (orgs.) *Contemporary composers on contemporary music*, p. 195-208. Nova York: Holt, Rinehart, and Winston, 1967.

WEN-CHUNG, Chou. Open rather than bounded. *Perspectives of New Music*, v. 5, n. 1 (1966), p. 1-6. Disponível em: http://www.jstor.org/stable/832383

WESTERKAMP, Hildegard. *Talking Rain*. [Obra eletroacústica, CD Harangue 1]. North Vancouver: Earsay Productions, 1998. Disponível em: http://www.earsay.com/soundshop/earsay/CDs/h1.html

WILSON, Scott. Spatial swarm granulation. In: *Proceedings of ICMC 2009* – International Computer Music Conference. SARC, ICMA, Belfast, 2008. Disponível em: http://eprints.bham.ac.uk/237/1/cr1690.pdf

WINDSOR, W. Luke. *A perceptual approach to the description an analysis of acousmatic music*. Londres, 1995. Tese (Doutorado) City University.

YOUNG, John. Reflections on sound image design in electroacoustic music. *Organised Sound*, v.12, n. 1 (2007), p.25-33.

# Composição textural algorítmica via sistemas dinâmicos

JÔNATAS MANZOLLI e ADOLFO MAIA Jr.

Evolving metaphors, in my opinion, is what artists do.

- Brian Eno<sup>1</sup>

ste capítulo introduz uma metodologia para a composição musical utilizando-se de modelos formais. Espera-se que esta abordagem sirva não só para compositores, mas para todos os interessados em música contemporânea e tecnologias recentes com as quais é possível desenvolver heurísticas e formalizações para gerar, organizar, moldar e ordenar estruturas musicais. Descreve-se aqui, sem a pretensão de esgotar o tema, a criação de dois modelos formais e os algoritmos computacionais *GYorGY* e *GYorGYscape*.

O texto que inspira o termo modelo formal é o livro Formalized Music (XENAKIS, 1971), obra fundamental que deve sempre ser consultada. Não desenvolvemos aqui nenhum estudo de modelos estocásticos, como fez Xenakis. O nosso enfoque vincula-se a outro ponto de vista: o compositor cria modelos determinísticos para comportamentos que evoluem no tempo ou, como denominaremos mais tarde, processos composicionais iterativos.

No século XX há um rico desenvolvimento de modelos formais aplicados à composição. Pioneiros como Lejaren Hiller (1959), Iannis Xenakis (1971; 2001), Milton Babitt (2003) e

\_

<sup>&</sup>lt;sup>1</sup> Esta epígrafe é citada de uma palestra proferida em São Francisco, EUA, em 8 de junho de 1996.

outros que trouxeram contribuições originais e importantes para este campo. Há que se notar que a música contemporânea beneficiou-se de modelos formais ou quase formais, mesmo que a tecnologia não estivesse representada objetivamente nos modelos. As proposições do Serialismo Integral repousam sobre certo grau de formalismo (FORTE 1973). O Minimalismo apresenta exemplos de automatismo aplicado à composição (MERTENS 1983; REICH 2002).

Mesmo que o compositor não tenha utilizado diretamente o formalismo para compor é possível analisar o processo composicional transportando sua estrutura para um sistema formal e vice-versa. Partindo deste ponto de vista, propomos aqui um exercício composicional abstraindo algumas propriedades do automatismo da escritura textural de György Ligeti. Alguns dos processos por ele utilizados para gerar padrões automáticos de alturas e ritmos evoluindo no tempo exemplificam o conceito de *iteração* na construção de modelos formais. Os sistemas dinâmicos que apresentamos a seguir são definidos como processos iterativos: *estruturas complexas são geradas utilizando-se repetição de padrões mais simples.* 

As próximas seções foram escritas também como processo iterativo. Partimos do que consideramos elementar para o entendimento de sistemas formais e dinâmicos para gradualmente evoluir os conceitos até apresentar sistemas texturais algorítmicos. O objetivo é mostrar que modelos composicionais, mesmo sendo formais, são vinculados a um ponto de vista poético. Pois a criação de modelos para composição algorítmica é uma atividade interdisciplinar na qual o compositor busca um equilíbrio pessoal entre objetividade e subjetividade. Notadamente, trata-se da transposição de um conceito de natureza subjetiva para um sistema de natureza objetiva ou formal. É preciso imaginar um conjunto de regras elementares que possam reproduzir uma evolução temporal de objetos sonoros.

# Pontos de partida

Os compositores contemporâneos criaram modelos aritméticos e combinatórios (BABBITT 2003) ou utilizaram a

seção áurea (ver análise das obras de Bartók em LENDVAI 1971). Curtis Roads (1996) apresenta modelos denominados de gramáticas generativas para definir sistemas de composição. Outro exemplo interessante vem da peça para piano *Klavierstücke XI* (1952) de Stockhausen que utiliza regras formais derivadas de matrizes (TRUELOVE, 1998). Há também trabalhos sobre os processos utilizados na série Klavierstücke. O chamado "Problema de Stockhausen" é estudado na área de combinatória (READ e YEN 1996). Para uma ampliação do estudo de modelos composicionais vide o capítulo 4 (BARREIRO e KELLER 2010).

#### Processo e sistema formal

Neste capítulo relacionamos sistema formal com a noção de processo musical para construir um modelo para composição textural algorítmica. Recomendamos uma leitura adicional sobre planejamento textural (ALVES 2005).

Modelos composicionais podem ser utilizados para gerar evolução automática de padrões no tempo como no processo de deslocamento de fase (*phase shifting*) criado pelo compositor minimalista Steve Reich (REICH 2002). Em *Clapping Music*, composta em 1972, Reich simplesmente superpôs duas figuras rítmicas, deslocando uma delas de um número crescente de pulsos ao longo da peça.

A noção de processo está presente em contextos musicais da atualidade como o minimalismo (MERTENS 1983) e o espectralismo (MOSCOVICH 1997; MURAIL 1980). Intuitivamente, processos musicais são transformações temporais aplicadas ao material utilizado na composição. A composição minimalista utiliza processos graduais que são construídos com a repetição ou iteração de padrões que, com pequenas variações, gradativamente conduzem a estrutura da obra de uma configuração a outra. O exemplo típico é a técnica desenvolvida por Reich denominada de deslocamento de fase mencionada acima e estudada em Cohn (1992). O conceito de sistemas dinâmicos, definido a seguir, engloba a noção de processo gradual. É possível criar uma evolução temporal gradativa, comportamentos cíclicos e até mesmo caóticos que

crescem de intensidade, acelerando, desestabilizando o fluxo sonoro.

Na obra *Partiels* para 18 instrumentos (1975), Grisey utiliza o conteúdo espectral da nota mi grave do trombone para construir um processo espectral. Alguns compositores espectrais conduzem o material sonoro a configurações texturais densas através de processos como adensamento, saturação e aglutinação. Outros reduzem o material produzindo texturas rarefeitas por diluição, alargamento, dissipação e erosão. Há processos rítmicos com a superposição de regularidades e irregularidades de pulso. Há também processos que transformam estruturas verticais harmônicas em inarmônicas ou vice-versa utilizando filtragem digital de um sinal musical (CERVINI, 2008).

# Poema Sinfônico como modelo

130

O modelo composicional que desenvolvemos neste capítulo está relacionado com o sistema mecânico criado pelo compositor húngaro György Ligeti no *Poema Sinfônico para 100 Metrônomos*, de 1962. Nessa obra, 100 metrônomos com diferentes pulsos geram uma textura complexa de tic-tacs (CONE 1977). O sistema evolui até a parada gradual de cada um dos metrônomos, o último tic ou tac no fim da peça. Este procedimento cria uma jornada consistente que parte da ordem absoluta do pulso para atingir um caos sonoro massivo e complexo antes de retornar ao silêncio puro e simples.

Transcrevemos abaixo um excerto da partitura, que na verdade é um texto em forma de bula<sup>2</sup>.

<sup>-</sup>

<sup>&</sup>lt;sup>2</sup> "Poeme Symphonique" may be performed in two versions:

<sup>1)</sup> All metronomes are wound equally tightly. In this version the chosen metronome numbers (oscillation speeds) wholly determine the time it will take for the several metronomes to run down: those which swing faster will run down faster, the others more slowly. 2) The several metronomes of a group are wound unequally: the first of the 10 metronomes the tightest, the second a little less, the tenth, the least tightly. Care must be taken, however, that the winding and the regulation of the speeds of the several metronomes are carried out completely independently of each other. Thus the metronome in each

[..] "Poema Sinfônico" pode ser executado em duas versões: 1) Todos os metrônomos são calibrados dando-se a mesma corda. Nesta versão o valor do andamento de cada metrônomo (velocidades de oscilação) determina totalmente o tempo gasto para os vários metrônomos pararem: aqueles que oscilarem mais rápido pararão mais rápido e os outros mais lentamente. 2) Vários metrônomos de um grupo são calibrados diferentemente: o primeiro de um grupo de 10 metrônomos terá mais corda, o segundo um pouco menos até o décimo com a menor corda de todos. Deve-se tomar cuidado, entretanto, para que a corda e a regulagem das velocidades dos vários metrônomos sejam feitas de forma independente uma das outras. Desta forma, o metrônomo em cada grupo que tenha recebido a menor quantidade de corda não precisa ser necessariamente o mais rápido ou o mais lento em suas oscilações. [..](LIGETI, 1962).

O Poema Sinfônico para 100 Metrônomos é nosso ponto de partida e implementamos o sistema GYorGY (caso unidimensional) e GYorGYscape (caso multidimensional) em linguagem Pure Data (PUCKETT 1996) como uma forma de exemplificar a técnica composicional discutida neste capitulo. Os princípios composicionais de Ligeti que motivaram a criação desses dois sistemas foram analisados em Manzolli e Luvizotto (2009).

Vários autores apontam para duas características importantes da escritura ligetiana: o automatismo algorítmico e a superposição de camadas. Em obras como "Continuum" (1968), a técnica de interpretação do cravo foi habilmente utilizada por Ligeti. Com a repetição de padrões mecânicos no teclado desse instrumento, ele conecta a sonoridade beliscada do cravo às nuances de um contínuo de alturas. Este método é descrito como "composição baseada em padrões mecânicos" (CLENDINNING 1993) que gera uma estrutura de redes (netstructures) como teias de aranha (ROIG-FRANCOLI 1995).

Ligeti interessou-se também pela Geometria Fractal de Benoit Mandelbrot (1982). Ao descrever seu *Piano Concerto* 

group, which has been most lightly wound must not be the fastest or the slowest in its oscillation.

(1985-1988), Ligeti relaciona os fractais com sua obra: "[os fractais] são sempre novos, mas sempre os mesmos e isto é algo que me atrai muito e encontra-se na minha música" (apud MANZOLLI, 1993).

A partir destas ideias descrevemos a seguir os passos da nossa modelagem. Desde as definições formais, a transferência das ideias do modelo composicional para o processo algorítmico e, finalmente, a análise das texturas geradas com esse modelo.

# Definições formais

O pensamento relacionado à modelagem matemática leva a uma abstração: um modelo do fenômeno com a definição de objetos, variáveis, funções e relações apropriadas que delimitam a compreensão do fenômeno. Com esse modelo, que é um tipo de reducionismo filosófico, dizemos que "entendemos", pelo menos parcialmente, o que acontece no interior dos átomos, das estrelas, no núcleo das células ou no cérebro humano. Esta abstração é denominada aqui, de maneira bastante geral, de *modelo formal ou sistema formal*. Neste texto adotamos a seguinte definição:

Um **Sistema Formal** é um conjunto de objetos entre os quais podem ser definidas relações lógicas consistentes.

Ao observarmos a natureza verificamos que muitos sistemas físicos e biológicos se transformam ao longo do tempo, mudam de forma, de posição ou envelhecem. No fim do século XIX e ao longo do século XX este tipo de fenômeno físico inspirou os matemáticos a criarem uma nova área do conhecimento, denominada de *sistemas dinâmicos*. Ela é bastante explorada atualmente, com pesquisas interdisciplinares em genética, engenharia aeroespacial, sistemas biológicos, engenharia elétrica entre outras. A partir deste ponto de vista, definimos que:

Um **sistema dinâmico** é um sistema formal que possui algum tipo de evolução temporal, isto é, que muda seu estado

em relação a um parâmetro que representa a variação do tempo.

## Estados, atratores e conjuntos-limite

Um conceito básico sobre sistemas dinâmicos é o de estado. Intuitivamente falando, o estado de um sistema é uma coleção de informações que podemos observar num determinado instante do tempo. Os valores dos estados de um sistema são obtidos através de variáveis associadas às informações que queremos do sistema. Por exemplo, se o sistema for um copo de água podemos avaliar o estado físico da água pela temperatura, densidade, transparência, etc. Ao medirmos estas variáveis ao longo do tempo, analisamos a evolução do sistema. Uma sucinta formulação matemática de sistemas dinâmicos é apresentada no apêndice deste capítulo.

Definimos aqui um sistema composto com apenas um ponto. Esse ponto pode ocupar qualquer posição do semi-eixo positivo da reta real, isto é, o intervalo  $(0, \infty)$ . Vamos tomar como um estado deste sistema pontual a sua posição na reta real positiva, a qual é, por sua vez, o espaço de fase do sistema. Vamos agora definir uma dinâmica através da simples função  $f(x) = \sqrt{x}$ . Denotando  $x_0$  a posição (estado) inicial do sistemaponto definimos então os estados posteriores pela equação recursiva (iteração)

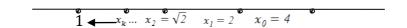
$$x_k = f(x_{k-1}) = \sqrt{x_{k-1}}$$
,  $k = 1, 2, 3, ..., L$ . (1)

A partir do estado inicial  $x_0$ , a órbita gerada é uma sequência de estados  $\{x_k\}_{k=0}^L$  e ela converge, isto é, se aproxima indefinidamente do estado  $\bar{x}=l$ . É fácil verificar isto usando uma calculadora e extraindo a raiz quadrada uma após outra. Dizemos que o estado  $\bar{x}=l$  é o *atrator* do sistema que, neste caso, é um simples ponto.

Existem sistemas mais complexos onde o atrator pode ser um conjunto de pontos. A dinâmica de um sistema pode apresentar comportamentos mais complexos do que uma simples atração, podendo também ser repulsão ou mesmo

ambas as características. Em geral tais conjuntos de pontos singulares são denominados *conjuntos limite*. Alguns sistemas podem apresentar, além de conjuntos limites, comportamento caótico. Tais sistemas dinâmicos não-lineares já foram utilizados como modelo composicional por Agostino Di Scipio (1990) e por Manzolli (1993) que estudou a aplicação de dinâmicas não-lineares em síntese sonora e composição em tempo real, respectivamente.

Voltando ao exemplo acima, vemos que a órbita  $\{x_k\}_{k=0}^L$  é finita com L+1 pontos, mas variando k indefinidamente podese obter uma órbita com um número infinito de pontos. O comportamento numérico convergente deste sistema dinâmico será utilizado a seguir na criação do nosso modelo composicional.



**Fig.1**: Sequência iterativa gerada pela equação (1) convergindo para o estado  $\bar{x} = 1$ 

#### GYorGY: modelo unidimensional

Uma audição detalhada do *Poema* leva-nos a considerar os seguintes comportamentos temporais: a) **dissipação gradual** de uma massa sonora criada pela superposição de pulsos, b) **desaceleração** desses pulsos causada pela perda de energia até a parada completa de cada um dos metrônomos e convergência ao silêncio. Os metrônomos do *Poema* não são sincronizados com o mesmo pulso e são acionados em momentos sucessivos. Este procedimento gera diferenças de fase entre os metrônomos e produz a imitação de padrões rítmicos. A superposição de figurações rítmicas com pequenas defasagens no tempo é um dos principais mecanismos da micropolifonia.

Para realização algorítmica do "Poema", podem ser utilizados sistemas formais diferentes, desde equações

diferenciais para o oscilador amortecido até o uso de funções iterativas levando ao comportamento caótico. A opção por um sistema ou outro, é justamente o que possibilita uma idiomática de modelos formais.

O modelo que criamos utiliza a Equação 1 apresentada na seção anterior. Lembramos que temos a sequência de números (estados) gerada iterativamente pela fórmula

$$x_k = f(x_{k-1}) = \sqrt{x_{k-1}} \ , \qquad k = 1, 2, 3, \dots, L.$$

e que esta sequência de números converge para o valor (estado)  $\bar{x}=I$ . Como este valor só é atingido no limite de um número infinito de passos, usamos o número de iterações L conveniente para a realização sonora desejada.

No nosso modelo composicional associamos a sequência convergente de valores numéricos (estados)  $\{x_0, x_1, x_2 \dots x_L\}$  a duração dos intervalos de tempo entre pulsos consecutivos de diferentes camadas sonoras. Utilizamos para cada camada a mesma equação, mas com parâmetros diferentes; como se fossem metrônomos com andamentos diferentes. As diferentes sequências numéricas equivalem ao comportamento dinâmico de cada metrônomo. Isto é o que denominamos de transporte ou mapeamento de um sistema formal para um modelo composicional algorítmico (ver glossário).

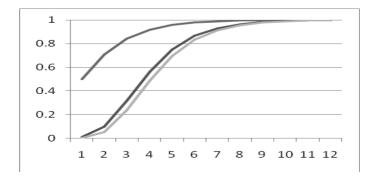
Ao analisarmos o efeito da sequência numérica no domínio sonoro, verificaremos que:

- se o valor inicial for  $x_0 > 1$ , os pulsos vão gradativamente desacelerar. Após L iterações os valores se aproximam do valor 1. Quanto maior for o número de iterações, mais o sistema se aproxima do estado de equilíbrio  $\bar{x} = 1$ .
- se o valor inicial for  $x_0 < 1$ , os pulsos vão gradativamente acelerar até um valor próximo de  $\bar{x} = 1$ .
- se o valor inicial for  $x_0 = 1$ , o pulso será constante e sempre igual ao estado de equilíbrio  $\bar{x} = 1$ .

A partir do comportamento numérico desse simples dinâmico. é possível imaginar estratégias sistema composicionais para gerar uma textura complexa superpondo camadas que convergem para o mesmo pulso (estado de equilíbrio), porém fora de fase umas com as outras. Cada camada sonora acelera, desacelera ou mantém um pulso regular. Diferente do *Poema* ligetiano, estes "metrônomos virtuais" convergem para um estado de equilíbrio que não é o silêncio, embora também seja possível definir parâmetros que possibilitem um estado final sem som. Na figura 2 mostramos um gráfico do comportamento de aceleração com valores numéricos iniciais  $x_0 \le 1$  que convergem gradativamente para a  $\bar{x} = 1$ .

# **GYorGYscape**: modelo multidimensional

O segundo sistema apresenta dois aspectos novos: *a)* cada camada sonora é associada a um vetor de parâmetros, dando origem a um controle multidimensional e b) existe interação entre os parâmetros que controlam as camadas. Se o processo sonoro resultante tiver com um grande número de camadas e os pulsos tiverem durações da ordem de milissegundos, estaremos simulando um processo de síntese granular (ver também BARREIRO e KELLER, 2010).



**Fig. 2:** O gráfico que descreve um acelerando gerado pela Eq. 1. No eixo vertical há os valores da raiz quadrada e no

eixo vertical os passos da função iterativa. Neste exemplo, foi utilizada uma sequência de 12 valores.

Recentemente, temos estudado síntese granular como um modelo do espaço tempo versus frequência de Gabor estendido para outras variáveis tais como espacialização, densidade de fluxo dos grãos sonoros e outras, formando também um espaço multidimensional com 6 variáveis independentes (FALCI e MAIA 2008) (ver glossário).

No *GYorGYscape*, além da duração e da frequência, há mais 7 outros parâmetros. Temos ao todo nove parâmetros para o k-ésimo estado do sistema: duração  $(t_k)$  dada em segundos, frequência  $(f_k)$  dada em hertz, intensidade (ou amplitude)  $(i_k)$ , valor inicial  $(x_k)$ , ataque  $(a_k)$ , decaimento  $(d_k)$ , sustentação (sustain)  $(s_k)$ , relaxamento (release)  $(r_k)$  e panorâmica (panning)  $(p_k)$ . Os cincos últimos parâmetros controlam um envelope ADSR que regula o perfil dinâmico do objeto sonoro e a sua projeção estereofônica. Para maiores esclarecimentos sobre ADSR vide o glossário e o Capítulo 2  $(LAZZARINI\ 2010)$ .

Como mencionado acima, se os valores da duração  $(t_k)$  forem abaixo de 100 milissegundos estaremos no domínio da Síntese Granular e os objetos sonoros se tornam grãos sonoros. Para durações maiores, estaremos gradativamente construindo estruturas com camadas harmônicas independentes, cada uma controlada pela interdependência entre os parâmetros frequência  $(f_k)$ , duração  $(t_k)$  e intensidade  $(i_k)$ . O parâmetro  $(x_k)$  controla um tipo de microdefasagem através da aceleração e desaceleração produzida pela Eq. 1, também definida com sincronia de fase em Barreiro e Keller (2010).

Para controlar este sistema utilizamos um operador matricial, ou seja, uma matriz quadrada **M**, 9x9, que gera a cada passo do sistema dinâmico um novo vetor de parâmetros. Formalmente temos a equação recursiva que representa os estados do sistema:

$$X_{k+l} = \mathbf{M}.X_k \tag{2}$$

A estrutura da matriz  $\mathbf{M} = \mathbf{m}_{ij}$  está diretamente ligada às estratégias composicionais. Optamos neste exemplo por usar uma matriz 9x9 formada por dois blocos diagonais:

$$\mathbf{M} = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \tag{3}$$

onde  ${\it A}$  é uma submatriz (bloco) 4x4 e  ${\it B}$  uma submatriz (bloco) 5x5. O bloco superior 4x4 controla a interação entre frequência, duração, intensidade e valor inicial. O bloco inferior 5x5 controla a interação entre o ataque, decaimento, sustentação, relaxamento de um envelope ADSR e a panorâmica. Os dois blocos na diagonal secundária são submatrizes nulas, isto é, contém apenas o valor zero. Optamos por utilizar a mesma matriz  ${\it M}$  para controlar todas as camadas da textura que geramos através do  ${\it GYorGYscape}$ . Determinamos também que os valores dos elementos da matriz satisfazem as desigualdades:

$$0 \le m_{ij} \le 1 \text{ para } i, j = 1, 2, ..., 9$$
 (4)

e para cada linha i, referente a uma das variáveis, a desigualdade

$$\sum_{j=1}^{9} m_{ij} \le 1 \tag{5}$$

Estas restrições sobre os elementos da matriz  $\mathbf{M}$ , garantem que o sistema descrito na Eq. 2, gere uma sequência convergente de vetores  $X_k$  para  $\mathbf{k}$  crescendo indefinidamente. SE a soma na Eq. (5) for igual a 1, a matriz  $\mathbf{M}$  é chamada Matriz Probabilidade de Transição e o processo descrito pela Eq. 2 é denominado Processo ou Cadeia de Markov. A prova matemática desta afirmação foge ao escopo deste capítulo, mas pode ser encontrada em qualquer bom livro de álgebra linear ou de processos de Markov (ANTON e BUSBY, 2006). Na verdade, para as nossas finalidades, as sequências de vetores são sempre finitas e, portanto, o problema da convergência não se aplica. No entanto os valores acima fazem com que os estados  $X_k$  se aproximem de um estado limite de equilíbrio  $\overline{\mathbf{X}}$ .

## 139

#### Análise das Texturas Geradas

A interface gráfica do *GYorGYscape* é apresentada na Figura 3. Na parte superior da interface estão controles deslizantes para frequência, duração, intensidade e valor inicial de cada camada. Na parte inferior, localizam-se controladores da reverberação e difusão sonora das camadas que são acionados apenas pelo usuário. A panorâmica e os quatro parâmetros do envelope ADSR são controlados por controles deslizantes individuais e os seus valores numéricos são associados simultaneamente a todas as camadas.

Do ponto de vista do controle matricial, 10 vetores, com 9 coordenadas cada um, são gerados automaticamente pelo sistema. Assim, o usuário e a sequência de matrizes controlam juntos o sistema em tempo real. As mudanças produzidas pelas matrizes ocorrem em todos os parâmetros simultaneamente enquanto que as mudanças produzidas pelo usuário ocorrem só no parâmetro desejado. O intervalo de tempo entre duas aplicações consecutivas da matriz de controle sobre os vetores é um parâmetro global acionado também pelo usuário ou por um arquivo texto definido a priori.

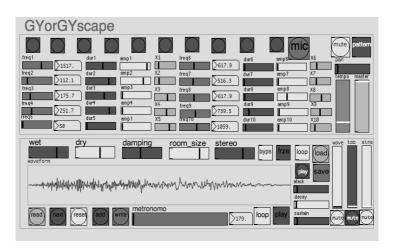
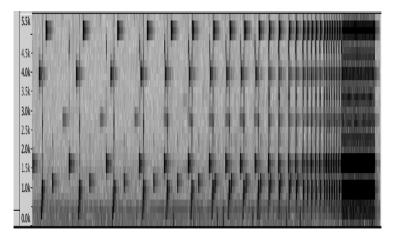


Fig. 3: Interface gráfica da implementação em Pure Data do GYorGYscape com 10 camadas texturais diferentes.

Analisamos a seguir, o espectrograma<sup>3</sup> de 3 texturas geradas pelos dois sistemas. A primeira textura, apresentada na Figura 4, foi gerada com o sistema *GYorGY*. Nela vê-se a evolução temporal de um processo de adensamento textural.

Uma análise visual do espectrograma mostra que a frequência de cada pulso permanece invariante e que o único parâmetro que se modifica com a evolução do sistema é o intervalo de tempo entre os pulsos controlado pelos estados do sistema dinâmico definido pela Eq. 1, ou seja, os valores dados pela sequência  $\{x_0,x_1,x_2...x_L\}$ 



**Fig. 4:** Textura gerada com o sistema *GYorGy*. Observa-se uma sequência horizontal de pulsos. Vê-se, da esquerda para direita, que o intervalo de tempo entre dos pulsos diminui até a acumulação. Os valores do eixo vertical estão no intervalo de 0 a 5 kHz.

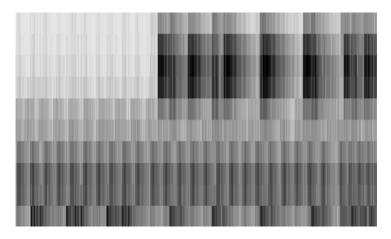
intenso.

<sup>&</sup>lt;sup>3</sup> Espectrograma é uma representação gráfica da evolução temporal dos parciais do espectro de um som. No eixo horizontal temos a variação do tempo, que da esquerda para direita, representa o início e o fim do som analisado. No eixo vertical temos a escala de frequência que, de baixo para cima, vai do grave ao agudo. Os tons de cinza representam a variação de intensidade, quanto mais escuro mais

**Fig.5:** Textura com as 4 camadas superiores e 2 inferiores síncronas e 4 camadas intermediárias assíncronas. Os valores do eixo vertical estão no intervalo de 0 a 10kHz.

A próxima textura apresentada na Figura 5, foi gerada com o sistema *GYorGYscape*, vê-se uma textura com 10 faixas espectrais que correspondem as 10 camadas, sendo 4 faixas superiores síncronas, 4 camadas intermediarias assíncronas e 2 camadas inferiores síncronas, novamente (ver glossário).

Na Figura 6 vê-se uma textura mais complexa com a acumulação de pulsos e com a variação de intensidades (descritos pelas variações de tons de cinza). Olhando as 10 camadas percebe-se que elas estão agrupadas em 4 subcamadas diferenciadas por padrões visuais diferentes. O espectrograma mostra a diferenciação assíncrona entre as camadas. No Figura 6 há maior variedade entre pulsos que pode ser vista pela variação de tons de cinza e a sua largura horizontal que é menos homogênea que na figura 5.



**Fig. 6:** Textura mais complexa com quatro subcamadas. Os valores do eixo vertical estão no intervalo de 0 a 10kHz.

142

Esses resultados ilustram o nosso modelo algorítmico para composição textural que é o resultado de trajetórias geradas com sistemas dinâmicos simples exemplificando o automatismo inspirado em Ligeti. Embora seja uma abordagem determinística, assim como é o *Poema*, o *GYorGY* e o *GYorGYscape* geram configurações texturais complexas. Enfim, o modelo mecânico de Ligeti descrito numa bula foi formalizado em termos de um conjunto de sistemas dinâmico simples independentes.

Este ambiente computacional permite a interação intérprete-sistema como um instrumento textural complexo. A entrada de dados pode ser feita através de interfaces gestuais. Para maiores detalhes sobre interfaces vide o capítulo 3 (WANDERLEY 2010).

#### Conclusão

Os modelos formais têm relação direta com o processamento algorítmico, o que possibilita a sua programação em linguagem computacional. O uso do computador engendra um campo praticamente infinito de possibilidades. Qualquer ideia tem o potencial de ser

representada na forma digital. Todavia trata-se de um paradigma que, de alguma forma, é reducionista e tem características próprias que não comportam todas nuances do fenômeno musical. Apesar dessa redução, o uso de modelos formais potencializa a criatividade, possibilita a expansão da inventividade do compositor que programa e desenvolve seus próprios processos. Ainda, o desdobramento de aplicações de sistemas dinâmicos em composição pode ser associado a sistemas evolutivos e bio-inspirados (MORONI *et al.*, 2000; FORNARI *et al.*, 2008) e a algoritmos genéticos que também geram texturas sonoras em tempo real (FELS e MANZOLLI, 2001).

Podemos fazer um paralelo entre o uso de modelos formais em composição e o estudo de contraponto. Nos dois há a necessidade de conciliar o processo criativo com um conjunto de regras. Toda regra é delimitadora e cabe ao compositor conduzir a estrutura da obra através do estilo escolhido. Da mesma forma, cada projeto de sistema formal possui uma visão específica de construção musical. Da mesma forma, neste capítulo optamos por explorar os sistemas dinâmicos como geradores de texturas. O modelo utilizado foi norteado por um ponto de vista: os processos automáticos da escritura micropolifônica de Ligeti.

A composição com algoritmos requer um tipo de *solfejo* de modelos formais. Não é uma exploração de parâmetros de um software até obter o resultado sonoro desejado. É uma busca de processos computacionais que viabilizam uma ideia musical embrionária. Xenakis (1971) criou modelos composicionais inspirados na física dos gases ideais; utilizou a distribuição estatística de Boltzmann para compor *Pithokrapta* no ano de 1956 e vários outros de seus modelos formais incluíram processos estocásticos, teoria de grupos e teoria de números (XENAKIS 2001).

Enfim, é relevante colocar a "mão na massa", construir modelos que possibilitam a criação de novos métodos composicionais e o controle de estruturas complexas que, muitas vezes, só é possível via uma escritura musical algorítmica.

## **Apêndice**

### Sistemas Dinâmicos: breve introdução formal

Há muitos exemplos de sistemas dinâmicos com leis de evolução diversas. Uma das mais usadas e conhecidas é através de *funções iterativas ou recursivas*. Primeiramente denotamos o estado de um sistema de N variáveis pela função vetorial  $\vec{E}(t) = (X_1(t), X_2(t), X_3(t), ..., X_N(t))$ .

Supomos que o sistema é observado no tempo  $t_0$  no estado  $\vec{E}_0 = \left(X_1(t_0), X_2(t_0), X_3(t_0), \dots, X_N(t_0)\right)$ . A dinâmica agora pode ser dada iterativamente por uma regra, relacionando o estado anterior com o posterior do sistema, do tipo:

$$\vec{E}_k = f(\vec{E}_{k-1})$$

com  $k=1,2,3,\ldots,K$  e onde K é o número total de iterações e f é uma função vetorial real denominada função iterativa ou função de iteração. O que temos é um sistema relacionando as variáveis em dois instantes de tempo consecutivos  $t_{k-1}$  e  $t_k$ , isto é:

$$\begin{split} &X_{1}(t_{k})\\ &=f_{1}(X_{1}(t_{k-1}),X_{2}(t_{k-1}),X_{3}(t_{k-1}),...,X_{N}(t_{k-1}))\\ &X_{2}(t_{k})\\ &=f_{2}(X_{1}(t_{k-1}),X_{2}(t_{k-1}),X_{3}(t_{k-1}),...,X_{N}(t_{k-1}))\\ &X_{3}(t_{k})\\ &=f_{3}(X_{1}(t_{k-1}),X_{2}(t_{k-1}),X_{3}(t_{k-1}),...,X_{N}(t_{k-1}))\\ &X_{N}(t_{k})=\\ &f_{N}(X_{1}(t_{k-1}),X_{2}(t_{k-1}),X_{3}(t_{k-1}),...,X_{N}(t_{k-1})). \end{split}$$

onde  $k=1,2,3,\ldots,K$  e as  $f_1,f_2,\ldots,f_N$  são as funções componentes da função vetorial f. Tipicamente, o sistema acima é um conjunto de leis de recorrência. Observe que o parâmetro k faz o papel do tempo, medido de uma maneira discreta. O conjunto de todos os valores possíveis das variáveis

 $\{X_k(t)\},\ k=1,2,3,\ldots,N,$  é denominado *Espaço de Fase* do Sistema (com N dimensões) ou *Espaço dos Estados* e o vetor  $\vec{X}_k=(X_1(t_k),X_2(t_k),X_3(t_k),\ldots,X_N(t_k))$  denomina-se *Vetor de Estado*.

Dado o vetor inicial  $\vec{E}_0 = (X_1(t_0), X_2(t_0), X_3(t_0), \dots, X_N(t_0))$ , para  $k = 0,1,2,3,\dots$ , K, este vetor de estado descreve um conjunto de K pontos no Espaço de Fase denominado de  $\acute{o}rbita$ . Mudando o estado inicial, muda-se também, em geral, a  $\acute{o}rbita$ , a não ser que se vá para outro estado da mesma  $\acute{o}rbita$ . Sistemas com duas ou mais variáveis podem gerar  $\acute{o}rbita$ s bastante complexas.

É importante notar que as funções componentes  $f_1, f_2, ..., f_N$  não precisam necessariamente ter uma fórmula. Por exemplo, elas podem ser definidas através de regras de associação, como é a própria definição de função que pode representar formalmente regras de composição, como no sistema GYorGY, onde utilizamos a mesma função para cada camada. Em vários sistemas dinâmicos utilizados na prática, as funções iterativas podem ser definidas algoritmicamente ou através de regras adequadas.

Dentre os sistemas dinâmicos que têm funções iterativas definidas por fórmulas matemáticas, os tipos mais simples são os *sistemas dinâmicos lineares*. Neste caso as funções componentes  $f_1, f_2, ..., f_N$  são funções lineares permitindo escrever o conjunto de equações iterativas simplesmente como uma equação matricial do tipo

$$\vec{X}_{k} = M\vec{X}_{k-1}$$
,  $k = 1,2,3,...,K$ ,

onde, para cada k o vetor de estado  $\vec{X}_k$  é uma matriz coluna  $N \times 1$  e  $M = (M_{ij})$ , com  $i,j = 1,2,\dots,N$ , é uma matriz  $N \times N$ . Daí, cada função componente do vetor de estado  $\vec{X}_k$  é uma combinação linear das funções componentes do vetor de estado  $\vec{X}_{k-1}$  precedente, ou seja,

$$\begin{split} X_n(t_k) &= M_{n1} X_1(t_{k-1}) + M_{n2} X_2(t_{k-1}) + \cdots \\ &+ M_{nN} X_N(t_{k-1}), \end{split}$$

para n = 1, 2, 3, ..., N.

O sistema GYorGYscape é um caso particular deste formalismo linear. Quando um sistema não obedece a uma dinâmica linear ele é denominado não-linear. Por exemplo, o sistema mostrado acima, com dinâmica gerada pela função raiz quadrada (vide Eq. 1), é não-linear, pois não satisfaz o princípio de superposição (ver glossário). O universo dos sistemas nãolineares é muito mais amplo e complexo do que sua contraparte entanto há grandes esforcos computacionais de aproximação do comportamento sistemas não-lineares utilizando-se aproximações lineares. Com o advento do computador, métodos numéricos e vários algoritmos foram desenvolvidos para calcular a evolução de sistemas dinâmicos com grande precisão.

### Referências

ALVES, José Orlando. *Invariâncias e disposições texturais: do planejamento composicional a reflexão sobre o processo criativo.* Campinas, 2005. Tese (Doutorado) – Instituto de Artes, Unicamp.

ANTON, Howard; BUSBY, Robert C. Álgebra linear contemporânea. Porto Alegre: Bookman, 2006.

BABITT, Milton. *The collected essays of Milton Babbitt.* Princeton: Princeton University Press, 2003.

BARREIRO, Daniel Luís; KELLER, Damián. Composição com modelos sonoros: fundamentos e aplicações eletroacústicas. In: KELLER, Damián; BUDASZ, Rogério (orgs.), *Criação musical e tecnologias: teoria e prática interdisciplinar*, p. 66-95, Goiânia: ANPPOM, 2010, Goiânia.

CERVINI, Lúcia. *Continuum, processo e escuta em "Territoires de l'Oubli": concepção de uma interpretação*. Campinas, 2008. Tese (Doutorado) – Instituto de Artes, Unicamp.

CLENDINNING, Jane Piper. The pattern-meccanico compositions of Gyorgy Ligeti, *Perspectives of New Music*, v. 31 (1993), p. 193-234.

COHN, Richard. Transpositional combination of beat-class sets in Steve Reich's phase-shifting music. *Perspectives of New Music*, v. 30, n. 2 (1992), p. 146-177.

147

CONE, Edward T. 1977. One hundred metronomes. *The American Scholar*, v. 46, no. 4 (Autumn 1977), p. 443-459.

DISCIPIO, Agostino. Composition by exploration of nonlinear dynamical systems. In: *Proceedings of ICMC 1990* – International Computer Music Conference, p. 324-327. San Francisco: International Computer Music Association, 1990.

FELS, Sidney; MANZOLLI, Jônatas. Interactive, evolutionary textured sound composition. In: *6th Eurographics Workshop on Multimedia*, pp. 153-164, 2001.

FORNARI, José; MAIA Jr, Adolfo; MANZOLLI, Jônatas. Soundscape design through evolutionary engines. *Journal of the Brazilian Computer Society*, v. 14, n. 3, p. 51-64, 2008.

FORTE, Allen. *The structure of atonal music.* New Haven: Yale University Press, 1973.

GRISEY, Gérard. *Partiels* para 18 instrumentos. Milan: Ricordi, 1975.

HILLER, Lejaren Arthur. Experimental music: composition with an electronic computer. Nova York: McGraw-Hill, 1959.

LAZZARINI, Victor. Técnicas fundamentais para instrumentos computacionais de síntese e processamento de áudio. In: KELLER, Damián; BUDASZ, Rogério (orgs.), *Criação musical e tecnologias: teoria e prática interdisciplinar*, p. 39-65, Goiânia: ANPPOM, 2010.

LENDVAI, Erno. *Béla Bartók: an analysis of his music*. London: Kahn & Averill, 1971.

LIGETI, Gyorgy. *Poema Sinfônico para 100 Metrônomos* 1962. Mainz: Schott Music.

LIGETI, Gyorgy. *Continuum* para cravo, 1968. Mainz: Schott Music.

LIGETI, Gyorgy. *Piano Concerto*, 1985-1988. Mainz: Schott Music.

MANDELBROT, Benoit. *The fractal geometry of nature*. Nova York: W. H. Freeman, 1982.

MANZOLLI, Jônatas. *Non-linear dynamics and fractals as a model for sound synthesis and real time composition*. Nottingham, Inglaterra, 1993. Tese (Doutorado) – University of Nottingham.

MANZOLLI, Jônatas; LUVIZOTTO, André Luiz. Análise de "Ramifications" de Ligeti utilizando Transformada Wavelet. In: *Anais do XIX Congresso da ANPPOM*, Curitiba, 2009.

MERTENS, Wim. *The American minimal music*. Londres: Kahn & Averill, Londres. 1998.

MORONI, Artemis; MANZOLLI, Jônatas; ZUBEN, Fernando von; GUDWIN, Roberto. Vox Populi: an interactive evolutionary system for algorithmic music composition. *Leonardo Music Journal*, v. 10 (2000), p. 49-54.

MOSCOVICH, Viviana. French spectral music: an introduction. *Tempo*, v. 200 (1997), p. 21-27.

MURAIL, Tristan. La révolution des sons complexes. In: *Darmstadt Beiträge zur Neuen Musik Ferienkurse*, p. 77-92. Mainz, London, New York: Schott, 1980.

PUCKETTE, Miller. Pure Data. In: *Proceedings of ICMC 1996* – International Computer Music Conference, p. 269-272. San Francisco: International Computer Music Association, 1996.

READ, Ronald C.; YEN, Lily. A note on the Stockhausen problem. *Journal of Combinatorial Theory*, series A, v. 76, n. 1 (1996), p. 1-10.

REICH, Steve. *Writings on music*. New York: Oxford University Press Inc, 2002.

148 ROADS, C. *The computer music tutorial*. Cambridge, Mass: MIT Press, 1996.

ROIG-FRANCOLI, Miguel. Harmonic and formal processes in Ligeti's net-structure compositions. *Music Theory Spectrum*, v. 17 (1995), p. 242-267.

TRUELOVE, Stephen. The translation of rhythm into pitch in Stockhausen's "Klavierstück XI". *Perspectives of New Music*, v. 36, n. 1 (1998), p. 189-220.

WANDERLEY, Marcelo M. Projeto e utilização de instrumentos musicais digitais. In: KELLER, Damián; BUDASZ, Rogério (orgs.), *Criação musical e tecnologias: teoria e prática interdisciplinar*, p. 39-65, Goiânia: ANPPOM. 2010.

XENAKIS, Iannis. Formalized music: thought and mathematics in composition. 2 ed. Hillsdale, NY: Pendragon Press, 2001.

XENAKIS, Iannis. *Formalized music: thought and mathematics in composition.* Bloomington: Indiana University Press, 1971.

## Da interação entre sons instrumentais e eletrônicos

RODOLFO COELHO DE SOUZA

o ponto de vista do ouvinte-receptor não há propriedades intrínsecas dos sons que permitam fazer uma separação absoluta e inquestionável entre sons instrumentais e sons gerados eletronicamente,¹ ainda que em nosso imaginário tal separação pudesse parecer evidente. Uma comprovação corriqueira dessa afirmação acontece ao ensinar música erudita do século vinte. Quando apresentamos a estudantes iniciantes, fragmentos do *Silicon Valley Breakdown* de David Jaffe e do *The Banshee* de Henry Cowell, alguns alunos classificam, erroneamente, a primeira peça como instrumental e a segunda como música eletrônica.

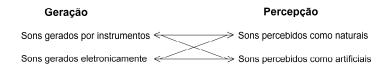
Na ausência da identificação visual da fonte sonora, recorremos à nossa memória sonora, que pode nos induzir a interpretações incorretas das fontes sonoras instrumentais e eletrônicas. Daremos a seguir alguns exemplos de como um efeito de *trompe l'oreille* (de enganar o ouvido) decorre, em última instância, de um efeito de *trompe l'oeil* (de enganar a visão). Lembremos ainda que, no dia a dia, ouvimos muito mais música através de dispositivos eletrônicos (isto é, alto-falantes de rádio, televisão, cinema, tocadores de CDs, fitas, MP3, etc.)

\_

<sup>&</sup>lt;sup>1</sup> Note-se que, como ponto de partida, a diferença entre sons instrumentais e eletrônicos é tomada inicialmente ao pé da letra: sons instrumentais são aqueles gerados por vibrações mecânicas de instrumentos acústicos, enquanto sons eletrônicos são gerados por circuitos eletrônicos, analógicos ou digitais, tornados audíveis por alto-falantes e amplificadores. A seguir essa separação quase tautológica é questionada no lado da percepção.

do que diretamente de instrumentos acústicos ao vivo. Mal nos damos conta de que em todos estes casos assumimos a representação da onda sonora gravada como um substituto perfeito do som do instrumento original. É portanto sintomático que nosso objeto de estudo desponte justamente na era dos videoclips, em que a audição fica condicionada pelo hábito de se ouvir música sempre acompanhada por imagens. Nos gêneros multimidiáticos o que frequentemente está em jogo é a representação das fontes sonoras. Ela nos remete ao problema da diegese sonora, estudado pela Teoria do Cinema (XAVIER, 2005, p.37).

O problema da interação entre sons instrumentais e eletrônicos tem portanto uma dupla articulação: o momento da produção postula, por motivos meramente técnicos, uma separação entre sons gerados por instrumentos e sons sintetizados, mas no momento da recepção essa separação pode encontrar-se mascarada por ambiguidades perceptivas. Ora, sabemos que o trabalho estético depende sempre de ambiguidades, isto é, de leituras polissêmicas (isto é, de significações múltiplas). Uma poética para a interação entre fontes de naturezas distintas pode servir-se dessas ambiguidades na recepção.



**Fig. 1:** Oposição entre sons naturais e artificiais na geração e percepção

Conclui-se então que a dialética da interação entre sons instrumentais e sons eletrônicos depende do reconhecimento de similaridades acústicas na ponta da percepção. Em ambas as colunas da Figura 1 há um contínuo de possibilidades. Os sons gerados por instrumentos podem ser idiomáticos, mas também podem ser pouco familiares, como os produzidos por técnicas instrumentais expandidas (os sons multifônicos de flauta, por

exemplo). Os sons gerados eletronicamente podem ser pouco familiares, tais como certos sons produzidos por algoritmos de também síntese granular, mas podem simular instrumentais com surpreendente perfeição. Entre esses extremos há um contínuo de possibilidades intermediárias, como, por exemplo, o caso dos sons instrumentais gravados e transformados eletronicamente, que poderiam estar situados em pontos intermediários entre os extremos de ambas as categorias. Portanto um campo de trabalho estético, idiomático a esse gênero misto, pode derivar da ambiguidade do contínuo de possibilidades nos campos da geração e da recepção.

# Uma perspectiva semiótica da interação: ícones, índices, símbolos e marcação

Uma reflexão sobre obras mistas através da tópica mais conhecida da Semiótica de Peirce, a da classificação dos signos em ícones, índices e símbolos (SANTAELLA, 2001 p.103) nos ajuda a perceber que as estratégias de interação entre sons eletroacústicos² e instrumentais correspondem a alternâncias do foco interpretativo do signo sonoro entre aquelas três categorias. Inicialmente consideremos que o aspecto de <u>ícone</u> puro, de primeiridade perceptiva, no qual se põe em relevo a <u>qualidade</u> intrínseca do som, é determinante para a própria existência do problema que estamos estudando. A criação eletroacústica encontra forte motivação na sedução da busca de sons nunca antes experimentados. Neste estágio não estamos preocupados com a relação entre os sons, mas sim com as

<sup>&</sup>lt;sup>2</sup> Note-se que utilizo aqui o termo "som eletroacústico" em vez de "som eletrônico". Ao lidar com problemas de significação relacionados à percepção, o termo "eletrônico" conota inevitavelmente a técnica de produção e uma gama de estilos, enquanto que "eletroacústico", apesar do prefixo "eletro", poderia ensejar interpretações mais neutras. Outrossim, consideraremos os sons como provenientes sempre de auto-falantes, sejam eles reconhecidos como instrumentais ou como eletrônicos.

152

Ora, se o aspecto icônico destaca as diferenças de qualidade entre os diversos sons, o aspecto indicial coloca em relevo a questão da fonte sonora. Nesta tipologia, o índice, para se configurar, depende de uma relação de contiguidade de fato entre significantes. O som do piano, por exemplo, é um <u>índice</u> instrumento piano porque em algum experimentamos uma relação de contiguidade entre o som do piano e sua fonte, o objeto piano, estabelecendo assim uma relação de causa e efeito: aprendemos que apertando a tecla do objeto piano fazemos surgir o som característico daquele instrumento. A partir da inversão dessa experiência na memória. continuaremos. toda vida. por decodificações indiciais que conectam sons semelhantes ao de pianos à existência implícita do objeto sígnico "piano" que, embora não presente, será hipotetizado como a fonte daqueles sons.

Devido à sua limitada aplicabilidade no âmbito de nosso estudo, deixaremos de lado a categoria do <u>símbolo</u>. Todavia deriva dele a compreensão de que a categoria "sons eletroacústicos" se coloca como *default*: todos os sons para os quais não se reconhece uma relação com uma fonte sonora real, recaem na categoria genérica "sons eletroacústicos". O conceito linguístico de marcação (*markedness*) nos ajuda a elaborar esse problema.

Marcação, como conceito teórico, pode ser definido de modo simples como uma atribuição de valor dada por uma diferença. [...] Dois termos de uma oposição assumirão

valores desiguais ou assimétricos, de "marcado" versus "não-marcado", com consequências para o significado de cada termo. [...] Por exemplo, quando usamos o termo *não-marcado* "vaca", não se distingue o sexo do animal; porém quando usamos o termo *marcado* "touro", especificamos "macho" e nesse contexto o uso do termo "vaca" teria mais chance de apontar para um sentido oposicional de "fêmea". Portanto "touro" é um termo *marcado* para a distinção de gênero, enquanto "vaca" é *não-marcado* (HATTEN, 1994, p. 34).

Podemos esquematizar a relação de marcação entre sons eletroacústicos e instrumentais através do gráfico da Figura 2, postulando a inclusão da categoria "sons instrumentais" como categoria marcada dentro da categoria não-marcada "sons eletroacústicos". De alguma maneira essa afirmação parece opor-se à nossa intuição, uma vez que não costumamos pensar nos sons instrumentais como abrangidos pelos eletroacústicos (mas também não pensamos nos touros como abrangidos pelas vacas...). De fato, o poder heurístico do conceito de marcação reside em sua capacidade de revelar processos de significação que parecem contrariar o senso comum. Para entender melhor esta conceituação, consideremos o exemplo de uma gravação da peça *Pluton* de Philippe Manoury, cujos sons foram todos gerados por um piano, às vezes tocado convencionalmente, outras vezes tocado com técnicas estendidas Р outras ainda com seus transformados eletronicamente até o limite de não serem mais reconhecidos como sendo sons de piano. Nesse universo, eletroacústico no sentido abrangente, certas características específicas dos sons de piano nos permitem distingui-los como pertencentes à categoria marcada "som instrumental de piano", embora de fato todos os sons sejam, em última instância, sons de piano.

Note-se que, na Figura 2, as bordas do campo "sons instrumentais marcados" foram desenhadas com linhas pontilhadas para representar a permeabilidade dessa fronteira às interpretações ambíguas, como já mencionado anteriormente.

Fig. 2: Marcação como princípio distintivo da interação

### De uma música baseada em notas à outra baseada em sons

Embora seja uma forma de expressão característica do pós-guerra, a música eletroacústica absorve algumas posturas típicas do modernismo da primeira metade do século vinte. Ela é obviamente herdeira do ruidismo dos futuristas italianos, de Russolo e Marinetti, e do projeto de uma música feita a partir de timbres, que encontra suas primeiras expressões em Varèse e em Schoenberg, com o princípio da *Klangfarbenmelodie*.

Entre as afinidades com o modernismo identificamos também o espírito iconoclasta de pregar a negação do passado. Tal como a "liberação da dissonância" havia sido essencial para atonalismo da Segunda Escola de Viena, a música eletroacústica, especialmente na vertente da música concreta francesa, pregou uma revolução sistemática dos parâmetros tradicionais da música. O uso de pulsos e ritmos recorrentes pareciam proibidos. O conceito de "nota" tornou-se obsoleto, na medida em que implicava em frequências fundamentais identificáveis e na estabilidade tímbrica de uma fonte instrumental reconhecível. Pelo contrário, postulava uma "escuta reduzida" apagaria que todos OS traços reconhecimento das fontes sonoras. Buscava-se uma poética estruturada na negação, isto é, não compor com notas (implicando não pensar a estruturação da música a partir das frequências), não compor com durações que fossem proporções geométricas do tempo (isto é, pulso, metro, compasso, ritmo) e não compor com timbres identificáveis. Restava tão somente a

transformação permanente de fragmentos sonoros, a utopia de uma música preocupada apenas com "o som em si".

Entretanto quase contemporaneamente às primeiras obras de música concreta e eletrônica gravadas em fita magnética. surgiam primeiras peças conjugando as instrumentos acústicos e sons gerados eletronicamente. Antokoletz (1992, p. 373) aponta a Música su due Dimensioni para flauta e tape de Bruno Maderna, composta em 1958, como a primeira peca desse gênero na história da música. Salientemos que essa obra magistral, ainda que talvez nãopropositalmente, divergia do projeto da música concreta francesa, entre outras coisas porque Maderna não abdicava de pensar em notas. Impregnada de pensamento serial na ordenação dos sons, tanto da fita magnética quanto da partitura para flauta, Maderna rejeitava (numa postura mais pósmoderna do que moderna) uma descontinuidade com as práticas do passado, ao mesmo tempo que procurava estratégias para expandir os procedimentos tradicionais. Uma dessas novas estratégias estava na sincronização entre flauta e fita utilizando princípios aleatórios para flexibilizar a rigidez da relação entre um tape imutável e uma parte instrumental que se renovava a cada nova interpretação.

Isto põe em evidência o mal-entendido de que uma música eletroacústica autêntica jamais poderia lidar com "notas," mas somente com "o som em si". Se essa separação fosse insuperável, peças mistas, em que instrumentos inevitavelmente tocarão notas enquanto alto-falantes reproduzirão sons eletroacústicos, jamais poderiam funcionar. Haveria uma imiscibilidade congênita entre os dois tipos de sons, como entre água e óleo, e o discurso das peças mistas sofreria de uma espécie peculiar de esquizofrenia acústica.

Aquela falácia nasceu de diferenças óbvias entre as duas práticas: compositores de música instrumental escrevem notas no papel, simbolizando sons que só se tornam audíveis pela mediação de um intérprete; compositores de música concreta trabalham diretamente sobre gravações de amostras sonoras. Compositores de música de meios mistos, porém, devem fazer as duas coisas. Se não houvesse algo em comum

entre as duas práticas, como seria possível almejar uma articulação entre elas?

## Tipologia espectral

Smalley (1986, p. 65-68), respeitado compositor tanto de música eletroacústica pura quanto de obras que misturam instrumentos acústicos com sons eletrônicos (ouça-se, por exemplo, *Clarinet Threads*, de 1985, para clarineta e tape), equaciona o problema acima exposto desenvolvendo uma tipologia espectral que postula a distinção entre notas (*notes*), nós (*nodes*) e ruídos (*noise*). A taxonomia de Smalley é uma simplificação da proposta anterior de Schaeffer (1966, p. 517) que reconhecia sete categorias (que citaremos mais adiante). Num texto mais recente, Smalley (1997, p. 119) simplifica ainda mais, descartando a tipologia do nó e concentrando-se nas outras duas. Há outras propostas de revisão de Schaeffer, como a Thoreson (2007), mas neste estudo optamos por usar as categorias do texto de Smalley de 1985, combinadas com detalhes adicionados em seu texto posterior.

A definição do conceito de ruído depende da oposição ao conceito de nota. Numa nota reconhecemos a predominância de uma frequência definida ou altura, enquanto o ruído só pode ser compreendido negativamente, pela inexistência de uma altura definida. Mas uma nota não se restringe a uma única frequência, é um conjunto de frequências, harmônicas em sua maioria. mas eventualmente também transientes componentes inarmônicos espalhados no espectro. Essa distribuição de frequências, também variável no tempo, está intimamente ligada à qualidade definida como o timbre do som. Como demonstra o gráfico da Figura 3, uma distribuição espectral pode ser perfeitamente harmônica (por exemplo, de uma nota gerada por síntese aditiva num computador), mas também pode ter um espectro substancialmente inarmônico. como ocorre, por exemplo, com o som dos sinos. Mesmo nesse caso ainda falamos de notas porque conseguimos estabelecer relações de intervalos entre os sons. Os carrilhões de sinos demonstram essa possibilidade. Mas quem já ouviu música tonal tocada num carrilhão conhece o estranhamento que nos

causa ouvir música pensada harmonicamente, materializada em sons com espectro inarmônico. Entretanto no espectro preponderantemente inarmônico da percussão de uma caixaclara com a esteira acionada é impossível identificar-se uma altura definida.



Fig. 3: Tipologia Espectral (adaptada de Smalley).

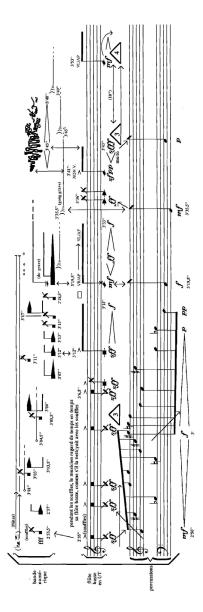
Observe-se que, à medida que a inarmonicidade cresce, a distinção entre notas e ruídos torna-se ambígua. No meio do caminho do contínuo entre esses dois extremos Smalley (1986, p. 67) estipula uma categoria intermediária, que ele denomina de <u>nó</u>, e que não chega a ser um ruído, mas "resiste à identificação de uma altura definida". O som de alguns instrumentos de percussão encaixam-se bem nesse paradigma, por exemplo, o som dos pratos suspensos e dos blocos de madeira. Nesses sons conseguimos reconhecer uma espécie de centro de gravidade da distribuição das frequências que nos permite fazer oposições de graves e agudos, embora seja impossível construir uma escala de sons afinados.

Do outro lado do contínuo situam-se os <u>Ruídos</u>. Observamos acima a importância do ruído na formação da linguagem da música eletroacústica. Isso foi possível porque a categoria dos ruídos abrange uma enorme diversidade de sons com diferentes qualidades. Smalley propõe dois critérios para classificar os ruídos: <u>granulação</u> e <u>saturação espectral</u>. Ruídos apresentam ademais qualidades subjetivas de rugosidade e granulosidade. São exemplos familiares de <u>ruído granular</u>, os sons de ondas marítimas e outras texturas de água, vento,

consoantes sem vogais e certos tipos respiração. O segundo critério, de saturação do espectro, que aliás não é conflitante com o primeiro, implica na percepção de um espectro compacto em que as frequências (inarmônicas) presentes estão próximas umas das outras e dentro de uma faixa suficientemente ampla para impedir a percepção de uma altura.

Podemos inquirir como interagem os sons instrumentais e eletrônicos sob o ponto de vista da tipologia espectral. Inicialmente lembremos que tanto as fontes instrumentais quanto as eletrônicas são capazes de produzir sons em qualquer posição do contínuo espectral acima descrito. É possível então organizar uma matriz de nove combinações em que linhas e colunas representam as três tipologias respectivamente instrumentais consignadas a sons eletrônicos. Interessante é imaginar que, ao longo do tempo, essas combinações podem variar e progredir de uma para outra. Obtemos assim um fundamento combinatório, sob o ponto de vista espectral, de uma linguagem idiomática da interação. Estamos, é claro, pensando no campo da produção dos sons, lembrando que, no lado da recepção, nosso propósito pode ser justamente jogar com a ambiguidade na percepção dessas nove categorias.

No fragmento reproduzido no Exemplo 1, extraído da peça *Parcours de l'Entité* (1991) de Flo Menezes, encontramos um exemplo desse tipo de poética de ambiguidades perceptivas entre sons das categorias de notas, nós e ruídos. A parte instrumental emprega uma flauta baixo e um conjunto de instrumentos de percussão metálicos (gongos de diversas origens e tamanhos, crotales, triângulos), associados na partitura às alturas das notas escritas. Os sons de percussão nessa obra tem sempre importantes componentes inarmônicos, situando-se em posições variáveis entre a tipologia de notas e nós. Os sons de flauta, por sua vez, podem ser de notas afinadas (como o Dó a 3'53") ou podem comportar tipologias variáveis entre nós e notas, dependendo da técnica expandida utilizada.



**Ex. 1:** Fragmento da partitura de *Parcours de l'Entité* de Flo Menezes.

No início do fragmento (2'55") a flauta é soprada quase sem definição de frequência, portanto representando uma tipologia a meio caminho entre o nó e o ruído. A 3'19,5" o flautista emite uma nota dó e ao mesmo tempo provoca um ruído com as chaves do instrumento. A 3'36" o compositor pede que se use um 'pizzicato de flauta' que consiste num golpe de língua, o que provoca outra qualidade de nó. A 3'41" o som da flauta se transforma de nó em nota e em seguida em ruído através do controle dos ataques, sustentação e impulso do crescendo.

Descrevemos até agora a variabilidade tipológica da parte instrumental. Na parte de fita, anotada na partitura com menos rigor, pois tal não se faz necessário para a execução, reconhecemos três fontes sonoras distintas: sons sintetizados eletronicamente com frequências definidas, sons gravados de flauta (executados com todos os tipos de técnica acima descritos para a parte instrumental) e sons de gongos. Cada um deles é mais ou menos transformado eletronicamente. Desse modo alguns podem ser classificados como notas (por exemplo, as notas identificadas como si bemol e dó, com timbre semelhante a flautas sintetizadas, no início do fragmento, ou o som de gongo grave a 3'19,5"), outros podem ser classificados como nós (como os diversos impulsos de sopro de flauta, transformados eletronicamente, entre 2'55" e 3'18,5" ou o som de gongo alterado a 3'35,5") e outros aproximam-se da tipologia de ruídos puros (como os eventos a 3'19,5 e 3'41").

Note-se que a interação entre as partes de fita e de instrumentos abrange uma significativa variedade de combinações tipológicas. Durante a performance ao vivo ocorrem diversas ilusões de fontes sonoras ambíguas entre o que é tocado ao vivo e o que está pré-gravado. As transformações dos sons gravados na fita muitas vezes são relativamente sutis, de forma a produzir uma espécie de performance virtual.

Percebemos assim que esta peça obedece com rigor à seguinte regra empírica:

Regra Prática Nº 1: Utilize na parte eletroacústica sons transformados de (ou emulando aspectos de) amostras dos mesmos instrumentos executados ao vivo, abrangendo uma ampla tipologia de notas, nós e ruídos.

Evidencia-se que a música de meios mistos recuperou, no âmbito da música eletroacústica, o desafio do uso de alguns parâmetros tradicionais. Ao mesmo tempo tornou mais evidente a ingenuidade do conceito de nota na música tradicional, uma vez que tal ideia representava uma redução radical das frequências presentes no som a apenas um componente fundamental. Na música eletroacústica esse tipo de redução simbólica perde relevância. Notas continuarão a aparecer em partituras, como no Exemplo 1, mas devemos ter em mente que ali elas são, muito mais do que na música tradicional, apenas traços mnemônicos de um fenômeno quase intangível, que é o universo da complexidade espectral dos sons utilizados na composição. Somente através da consciência da potencialidade das relações espectrais podemos mergulhar de fato nesse novo universo do "som em si," que não é baseado apenas numa escuta das relações formais entre os sons (como os intervalos de alturas), mas principalmente em suas qualidades intrínsecas.

## Tipologia morfológica

A tipologia espectral abordada acima preocupou-se com as frequências dos sons mas não cuidou do problema da evolução temporal do espectro no tempo. Sabe-se que as diversas frequências espectrais dos sons naturais apresentam intensidades variáveis. O perfil da variação dessas intensidades ao longo do tempo define a percepção da morfologia do som. Os dois tópicos considerados conjuntamente formam a chamada análise espectro-morfológica.

O texto inaugural de Schaeffer (1966, ampliado em 1977, p. 532) dava bastante atenção aos Critérios de Ataque para os quais reconhecia sete categorias. Num extraodinário tour de force, a empreitada taxonômica de Schaeffer culminava, no capítulo *Livro VI*, numa tabela abrangente e complexa, que

procurava exaurir as combinações de classificações a partir de sete critérios espectro-morfológicos: massa, dinâmica, timbre harmônico, perfil melódico, perfil de massa, grão e *allure* (ao pé da letra 'modo de andar', mas trata-se de um conceito schafferiano de difícil tradução que se refere ao processo de evolução temporal do som). A proposta de Schaeffer foi influente, principalmente por chamar a atenção para tipologias espectro-morfológicas pouco utilizadas até então, como as texturas granulares.

Seguindo a senda de Schaeffer, Smalley defendeu que a percepção da morfologia dos objetos sonoros é condicionada culturalmente pela nossa experiência com os sons instrumentais. Se, em outro contexto, poderíamos oferecer objeções à proposição de Smalley, em se tratando da interação entre sons instrumentais e eletroacústicos este é um postulado muito conveniente. O ponto essencial dessa proposta é a correlação de causa e efeito que vincula a produção do som instrumental à sua percepção. Elaborando este argumento Smalley afirma:

Durante a execução de uma nota, a energia dada [pelo intérprete] é transformada em alterações do conteúdo e da complexidade do espectro. Ao ouvirmos esta nota, invertemos a relação de causa e efeito, deduzindo fenômenos energéticos a partir de mudanças no espectro. O perfil dinâmico articula-se a mudanças espectrais: o conteúdo do espectro responde a forças dinâmicas, ou inversamente, forças dinâmicas são deduzidas a partir de alterações no espectro. Esta congruência auditiva entre perfis dinâmicos e espectrais, associados a fenômenos energéticos, é essencial para os procedimentos perceptuais rotineiros (SMALLEY 1986, p. 68).

O paradigma universal dos processos formais é a sequência "Início-Meio-Fim" que, em larga escala, abrange a forma de uma música e suas estruturas fraseológicas, e em pequena escala o desenvolvimento temporal de cada som. Na micro-escala da nota, este paradigma é usualmente representado nos aplicativos tecnológicos pelo diagrama ADSR

(traduzível em português por Ataque-Decaimento-Sustentação-Repouso, notando-se a subdivisão do "meio" em duas partes, D e S). Para mais detalhes sobre este tópico consultar, neste mesmo volume, o capítulo sobre técnicas fundamentais de áudio escrito por Victor Lazzarini.

Para Smalley (1986, p. 69 e 1997, p. 113) esse paradigma gera três arquétipos morfológicos: o ataque impulso, o ataque decaimento e a continuação gradual. No "ataque impulso", um início súbito é seguido imediatamente por um final abrupto, como no som do choque de duas pedras. Na "continuação gradual", o som é introduzido gradualmente, o que desloca nossa atenção para a fase intermediária de sustentação do som. As notas longas dos instrumentos de sopros seguem esse paradigma. Já o "ataque decaimento" comporta duas variantes: ataque decaimento fechado e ataque decaimento aberto. Em ambas as variantes o ataque é prolongado por uma ressonância, tal como numa nota de violão. As variantes fechada e aberta dão conta das diferenças na sustentação da ressonância em direção ao repouso. O som do xilofone, que tem um ataque abrupto e que decai rapidamente segundo uma curva exponencial breve, é classificado como "ataque decaimento fechado". O som do vibrafone, que também tem um ataque abrupto mas tem uma fase intermediária de sustentação mais prolongada antes da terminação exponencial, é chamado de "aberto".

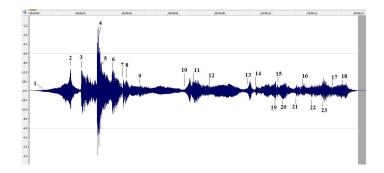
O arquétipo <u>ataque decaimento</u> é o que apresenta o melhor balanço perceptivo das três fases de ataque, sustentação e final, enquanto o <u>ataque impulso</u> concentra nossa atenção no ataque e a <u>continuação gradual</u> na sustentação. Note-se que esses paradigmas podem ser combinados de diversas maneiras de modo a gerar agrupamentos morfológicos complexos. Lembremos dos sons granulares que podem ser gerados, por exemplo, a partir da soma de múltiplos componentes do tipo <u>ataque impulso</u>.

Se em qualquer estilo musical as diferenças morfológicas dos sons são importantes na articulação do discurso musical, na música eletroacústica, especialmente no caso das interação de sons instrumentais e eletrônicos, elas podem desempenhar um papel mais fundamental ainda.

Analisemos um fragmento da partitura do *Synchronism Nº 9* de Mario Davidovsky para violino e tape (1988), que exibe uma generosa variedade de tipologias morfológicas (Ex. 2). Para facilitar a visualização da análise morfológica, fornecemos também um sonograma dessa passagem.



164



**Ex. 2:** Partitura e Sonograma de trecho do *Synchronism Nº 9* de Mario Davidovsky.

No primeiro evento dessa passagem o violino toca um intervalo de nona em duplo harmônico que principia como se fosse seguir o paradigma da <u>continuação gradual</u>. Entretanto o som que crescia gradualmente, é subitamente articulado por um

novo impulso, o evento 2, um acento forte que transforma a morfologia num <u>ataque decaimento fechado</u>. Este é um exemplo de que as morfologias podem se combinar em agrupamentos complexos, como já mencionamos acima. O evento 2 desencadeia uma série de <u>ataques decaimento</u>, alguns deles <u>abertos</u>, como os eventos 3 e 6 no violino, outros <u>fechados</u> como o evento 7 no violino, ou o evento 5, em pizzicato, um <u>ataque impulso</u> que, prolongado pelo tape, se converte num <u>decaimento aberto</u>, ou ainda o <u>ataque decaimento aberto</u> do evento 4 ouvido somente no tape. Constatamos assim a importância do jogo cruzado entre as tipologias morfológicas dos sons instrumentais e sintéticos para a eficácia do início desta passagem.

A continuação do fragmento é igualmente instrutiva. Davidovsky usa ali uma técnica de heterofonia<sup>3</sup> entre sons instrumentais e sintéticos: violino e tape tocarão juntos a mesma sequência de notas (fá#, ré, si, sol, si, conforme eventos 8 a 14) utilizando uma ampla variedade de espectromorfologias, algumas delas até imprevisíveis. O evento 8 é um ataque decaimento aberto que se prolonga no evento 9. No evento 10 ele faz um hábil uso de ataque gradual com decaimento rápido que se articula como continuação gradual através do prolongamento dado pelo evento 11 do violino (aliás, também novamente articulado, pelo ataque do evento 12) e que finalmente leva esta passagem à sua conclusão no evento 13, que surge com uma inesperada tipologia de continuação gradual, dada por um ataque exponencial com decaimento retrógrado, que encerra a passagem com outro belo exemplo de agrupamento morfológico complexo.

A parte final do exemplo (eventos 14 a 18 no violino e 14 a 23 no tape) exemplifica a dificuldade de uma notação eficaz na música mista para instrumento e tape. Já acontecera no evento 4 e novamente ocorre nesse trecho que a grafia da parte do tape nos induz a esperar um ruído ou um nó, embora

<sup>&</sup>lt;sup>3</sup> Heterofonia é uma técnica usual nas culturas com música monódica: uma mesma linha melódica é cantada e tocada por diversos instrumentos, com pequenas variações ornamentais, obtendo-se um efeito de complexidade sem haver polifonia de fato.

fundamentais identificáveis. Apenas os ritmos não são proporcionais, o que poderia ter sido resolvido com quiálteras. Provavelmente as notas dessa passagem foram geradas aleatoriamente, induzindo o compositor a representá-las como indeterminadas, embora de fato não soem assim. Esta passagem exemplifica um problema frequente na música eletroacústica. Mesmo quando não almejamos um controle definido de frequências, elas podem estar presentes, polarizando a recepção. Isto é, ruídos e nós podem incluir frequências identificáveis, soando como Notas. Como nosso ouvido está preparado para decodificar músicas com base em variações de frequências, a mera desatenção a esse parâmetro não faz surgir uma linguagem eletroacústica eficiente, ao contrário, pode conduzir ao desastre, com um discurso que não leva em conta o funcionamento da percepção. Não é o caso da obra de Davidovsky, onde, pelo contrário, parece até haver um excesso de atenção às notas, embora, na verdade, a real originalidade daquela peça resida na elaborada invenção espectro-morfológica da combinação entre sons instrumentais e eletrônicos.

de fato o que ouvimos é uma cadeia de ataques com

# Movimento virtual (categorias de gestos, texturas e figuras)

Pode-se dizer que não é idiomático da linguagem eletroacústica recorrer, como na música tonal, a expectativas criadas por dissonâncias e ciclos de dominantes para conferir direção ao discurso. Os efeitos de direcionalidade criados pela sintaxe eletroacústica são de outra natureza. Eles demandam que o ouvinte substitua os hábitos da escuta musical tradicional outra escuta sedimentada nos hábitos interpretação dos eventos sonoros do mundo físico. Por exemplo um fim de frase, equivalente a uma cadência tonal, pode ser modelado por uma mera diminuição de intensidade sonora, assim como, na experiência cotidiana, a dissipação de energia cinética leva à imobilidade do objeto.

Na música eletroacústica, a percepção de movimento pode resultar do contorno dinâmico de um <u>gesto</u>, do comportamento de uma <u>textura</u> (uma configuração formada

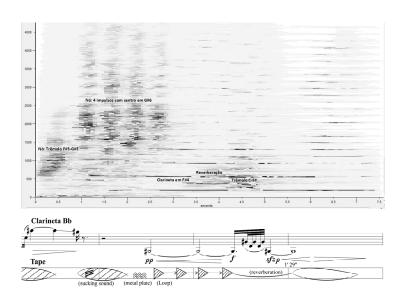
por um número grande de componentes semelhantes) ou do direcionamento produzido pela ordenação temporal dos componentes de uma figura (por exemplo a movimentação da tônica para a dominante implícita no perfil melódico do tema de uma sonata clássica). Tais categorias nos permitem elaborar certos aspectos fundamentais da sintaxe eletroacústica, tais como sentido de direcionalidade do fragmentação em agrupamentos percebidos como frases, por sua vez separados por articulações equivalentes a cadências. As relações espectro-morfológicas entre os objetos musicais, ainda que heterogênos, favorecem a formação de agrupamentos unificados por uma significação dinâmica, ou seja, pela sugestão de um movimento virtual. Em outras palavras, as relações descritas por esses conceitos atuam como uma espécie de cimento aglutinador de componentes heterogêneos, podendo conferir-lhes um sentido de direcionalidade que orienta a narrativa musical.

O conceito de **gesto** implica na coesão de componentes sonoros heterogêneos através da percepção de uma *Gestalt,* isto é, de uma <u>forma dinâmica unificadora</u> (Gillespie 1999 p.236). O uso desta técnica permite realizar a justaposição de objetos sonoros contrastantes, tal como a superposição de sons de uma clarineta ao vivo com ruídos pré-gravados, apenas alocados em contiguidade temporal, mas que, de alguma maneira, encontram um vínculo através de alguma similaridade icônica, às vezes até inesperada.

Isso acontece no exemplo 3, tomado da peça *Clariágua* (1999) de Rodolfo Coelho de Souza, para clarineta em si bemol e tape (que contém sons transformados de amostras de água), escolhido para ilustrar a tipologia do <u>gesto</u>. O começo do fragmento mostra o final da frase anterior, em que a clarineta faz um glissando de sol# para fá# (lembrar que se trata de instrumento transpositor), enquanto os sons gravados também desaparecem. Imediatamente, iniciando a frase subsequente, surge um som intenso, que oscila como um trêmolo e tem timbre que lembra, metaforicamente, "água sugada em placa reverberante". Segue-se outro som com quatro impulsos em laço, de timbre semelhante anterior, que reverbera a seguir, em direção ao final da frase, com um desenho dinâmico de crescendo-decrescendo. Simultaneamente ao bloco em laço,

entra a clarineta, recuperando a nota fá# com que terminara a frase anterior, mas duas oitavas abaixo, no registro peculiar do *chalumeau*. A clarineta desenha um crescendo que articula um breve trêmolo, desembocando novamente em fá#, com perfil dinâmico semelhante ao da reverberação que o acompanha.

A análise espectral de uma gravação desse fragmento revela certos detalhes interessantes. Fica claro que este gesto tem três partes distintas que cobrem regiões diferentes do espectro. A primeira é um nó (no tape, água sugada em placa reverberante) que oscila no registro médio entre fá#5 e sol#5. A segunda são os quatro *loops* em primeiro plano, com muitos harmônicos superiores e com a clarineta surgindo no fundo. Na terceira, invertem-se os planos e prevalece o perfil grave da clarineta em fá#4, com a reverberação da fita servindo de acompanhamento.



Ex. 3: Partitura e análise espectral de trecho de *Clariágua*, de Coelho de Souza

169

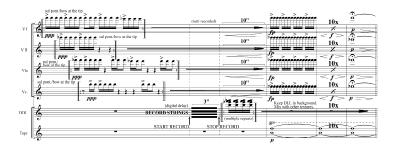
O espalhamento no espectro facilita a clara distinção entre os três objetos, mas ainda assim percebemos esta passagem como um único gesto porque há certos elementos que lhe conferem unidade gestáltica. A despeito dos sons gravados serem basicamente ruídos, eles contém frequências predominantes fá# e sol# que aparecem tanto na fita quanto na clarineta (aliás, usando de linkage technique com a frase anterior). O primeiro objeto do tape, classificável como nó, contém uma oscilação que repercute depois no trêmolo da clarineta. O segundo objeto do tape, outro nó, transforma o trêmolo em múltiplos impulsos e expande o espectro para as altas frequências. Os acentos, crescendos e decrescendos, tanto instrumentais quanto eletrônicos, articulam-se como uma sequência coerente e integrada, partindo de um impulso inicial e revitalizam-se em sucessivas ondas, até que a energia se dissipe no final. Em resumo: apesar de intrinsecamente diferentes entre si, os sons de ruídos de água e os sons de clarineta são utilizados na composição de modo a realçar suas semelhanças espectro-morfológicas, as quais permitem à nossa percepção aglutiná-los como pertencentes a um único gesto.

Neste caso o prolongamento da alternância entre as notas fá# e sol# produziu virtualmente um sentido de movimentação gestual, mas é necessário enfatizar que não se trata de um efeito de espacialização sonora, tal como o produzido por equipamentos de *surround sound*. Trata-se de um problema essencialmente cognitivo em que atribuímos significados espaciais a percepções que podem ser até monofônicas. As qualidades espectro-morfológicas da cadeia de objetos sonoros que compõe um gesto representam um movimento, em que nada necessariamente sai do lugar, mas em que um sentido de direcionalidade fica sugerido.

A segunda tipologia é a das **texturas**. Uma definição operacional que cobrisse todos os usos desta categoria é difícil. Mas podemos adotar a proposição de Ferneyhough (1995, p. 386) que define "texturas-classes como um certo tipo de fluxo sonoro formado por miríades de elementos de curta-duração, com características espectro-morfológicas que lhes conferem uma similaridade global, variável dentro de uma certa gama, com tendências de transformação processual, e que desenham configurações típicas, ainda que não uma figuração regular

identificável". Assim como os veios de uma prancha de madeira, uma textura sonora deve exibir, num certo trecho, um certo padrão médio de granulação, mas não um desenho perfeitamente repetitivo. Usualmente a distribuição desses elementos pode ser modelada por alguma regra estocástica, mas figurações determinísticas suficientemente complexas também podem encaixar-se neste modelo. Geralmente os componentes de uma textura mudam ao longo do tempo, produzindo um processo de evolução gradativa do conjunto que pode ser usado como um recurso efetivo para o prolongamento, o direcionamento ou a articulação das estruturas fraseológicas.

Cabe aqui ressaltar o que já foi dito antes: num gesto nossa atenção é absorvida pela envoltória global do objeto sonoro, pela superfície de seu impulso dinâmico, enquanto numa textura nossa atenção concentra-se nos detalhes da massa sonora, na contemplação de seus movimentos internos, nas relações microscópicas entre as partes do fluxo sonoro. Embora certamente não exista um gesto puro ou uma textura pura. É sempre possível focar nossa atenção nos componentes internos de um gesto, percebendo sua textura, assim como é possível focar a atenção na evolução da envoltória global de uma textura, ainda que lenta, percebendo sua gestualidade. Portanto, a dicotomia gesto versus textura é um processo dialético que depende do foco de atenção da nossa percepção. Todavia, geralmente é possível distinguir os trechos em que prevalece um sentido gestual, daqueles em que prevalece o aspecto textural. Reconheceremos imediatamente que essas categorias cumprem funções diferentes na configuração da fraseologia musical, e, mais ainda, que a alternância entre gestos e texturas é um recurso fundamental para se manter a variedade e o interesse do discurso na música eletroacústica.



Ex. 4: Partitura de excerto do Quarteto de Cordas No.1 de Montague

A partitura do Exemplo 4 ilustra a tipologia das texturas com uma passagem do *Quarteto no. 1* (1993) para cordas e sons eletrônicos, de Stephen Montague. Os instrumentos do quarteto criam inicialmente uma textura instrumental gerada pela sobreposição aleatória de módulos repetidos. Num certo momento, indicado na partitura, um operador de áudio inicia a tomada ao vivo de 3 segundos do som do quarteto. Essa amostra é armazenada na memória de um dispositivo de DDL - *digital delay line* (linha de retardo digital) e a seguir é reproduzida em laço fechado nos altofalantes, em múltiplas repetições, criando o efeito característico do dispositivo de retardo ao qual ainda se mistura, durante 10 segundos, a textura do quarteto ao vivo.

Obedecendo à recomendação da regra prática nº1, Montague cria um adensamento gradativo da textura através da acumulação estocástica de materiais com espectro-morfologia similar. Obtém assim uma perfeita simbiose entre os sons instrumentais ao vivo e os sons eletrônicos reproduzidos em retardo. Trata-se de uma técnica simples, mas efetiva. Após 10 segundos, o compositor modula bruscamente a textura instrumental e acrescenta sons eletrônicos de altura fixa, reavivando nosso interesse no caminhamento gradual da sonoridade.

A última tipologia, a das **figuras** nos remete à tradição dos motivos, frases, temas e períodos da música instrumental tradicional, na medida em que esta se baseia na ideia de <u>nota</u>,

ou seja, de sons musicais baseados em alturas e proporções temporais. Smalley não utiliza esse paradigma porque na música eletroacústica pura em que está interessado, figurações raramente ocorrem. É Ferneyhough (1987) que define esse paradigma no âmbito da música instrumental.

A percepção de uma figura depende da repetição, variada ou não, de uma certa configuração de notas. Por isso Xenakis insistia na diferença essencial entre figuras e texturas alertando que estas últimas dependem de "fenômenos macroscópicos em o que o resultado total da massa sonora é o que conta [...] Microsons e grãos elementares não tem importância na escala que escolhemos, somente grupos de grãos e as características destes grupos é que tem significado" (Xenakis 1992, p. 49-50). Entretanto na interação entre instrumentos e sons eletrônicos encontramos bordas difusas entre os domínios da figura e da textura, pois a linguagem instrumental carrega consigo a pesada tradição da figuração e os recursos eletrônicos tem mais afinidade com os processos estocásticos. Na interação entre eles o peso pode deslocar-se para um aspecto ou para outro. Por isso não é de estranhar que em todos os exemplos analisados acima se possa reconhecer uma forte participação de figuras na construção do discurso.

Vimos assim que a estruturação dos materiais sonoros através da alternância de três paradigmas sintáticos (gesto, textura e figura) permite que o discurso musical encontre, momento a momento, uma unidade orgânica, ainda que composto de materiais heterogêneos. Apesar das fontes de produção gerarem diferenças semânticas e perceptivas entre os sons instrumentais e eletrônicos, as relações sintáticas promovidas pelas três tipologias produzem significações locais de escala maior que unificam os sons.

## Aspectos de sintaxe e parataxe

A teoria dos gestos, texturas e figuras supõe que o discurso musical se organize sempre de modo sintático, isto é, por subordinações entre as partes. Este é um resquício da tradição analítica da música instrumental tradicional. Na modernidade, e especialmente na pós-modernidade, os

elementos do discurso nem sempre se articulam de modo sintático. Kendal (2006) analisa esse fenômeno na linguagem do *Poema Eletrônico* de Edgar Varèse, uma obra magistral composta ainda na infância da música eletroacústica. Naquela peça os materiais sonoros muitas vezes aparecem apenas justapostos, sem uma lógica evidente. Não há como justificar os encadeamentos por critérios sintáticos. Os sons demandam que nossa imaginação <u>crie</u> um sentido. Esse tipo de organização do discurso é chamada de <u>paratática</u>. Ela aparece no discurso da criança que narra uma sequência de eventos, sem obeceder a uma ordem aparente, justapondo fatos sem relação de causa e efeito, valendo-se de conjunções coordenativas.

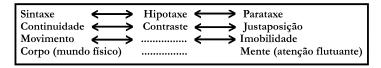


Fig. 4: Possibilidades de estruturação sintática do discurso

Note-se que, ao contrário das estruturas sintáticas (que no universo da música são as responsáveis pela criação de continuidade, sentido de direcionalidade e movimentação virtual), as estruturações paratáticas permitem à nossa mente flutuar à deriva. Aliás não se confunda esse tipo de suspensão com a estaticidade das texturas. Se uma textura carecer de direcionalidade em sua envoltória global, em compensação orientará nossa atenção para o movimento interno de seus componentes. A suspensão da parataxe é de outro tipo, pois atua sobre a articulação do discurso, que passa a ser fragmentário, carecendo de um vetor resultante que nos empurraria numa direção clara. Nem por isso a justaposição paratática é menos relevante ou frequente na linguagem musical, especialmente na interação entre sons instrumentais e eletrônicos, em que o potencial de contraste por mera justaposição pode ser muito eficaz. A Figura 4 esquematiza esse problema da oposição sintaxe-parataxe. Note-se que a meio caminho entre duas categorias encontram-se, as frequência, configurações intermediárias categorizadas como hipotaxes que também dependem de efeitos de contraste.



Ex. 5: Início de *Colorless Green Ideas Sleep Furiously* de Rodolfo Coelho de Souza

174

O Exemplo 5 reproduz a frase inicial de *Colorless Green Ideas Sleep Furiously* (1998), de Rodolfo Coelho de Souza, para piano e tape, na qual se alternam articulações sintáticas e paratáticas. No primeiro compasso o piano introduz um material temático, imediatamente respondido no tape por sons eletrônicos de alturas similares, criando um vínculo sintático entre eles. Entretanto nos compasso 3 aparece, no tape, um som estridente (identificados na partitura como "buzz") cuja relação espectro-morfológica com os demais materiais do fragmento é tênue. Do mesmo modo, a figuração pianística do compasso 4 parece fora de contexto, uma justaposição arbitrária. Ainda assim, devido a repetições ao longo da peça, atribuímos sentido a essa sequência de eventos, ainda que a sintaxe da frase nos pareça inicialmente frouxa, isto é, paratática.

# Contraste ou semelhança: o que é mais eficiente na interação?

A oposição entre parataxe e sintaxe acima descrita revela a existência de diferentes níveis de semelhanças e contrastes entre os materiais sonoros. Menezes defende que:

é comum encontrar-se a concepção errônea de que a interação deve basear-se exclusivamente na *fusão* entre a escrita instrumental e os dispositivos eletrônicos, uma vez que o *contraste* entre estas esferas sonoras é tão significante quanto os estados *fusionais*. Apesar de que a fusão possa ser vista como o mais importante ingrediente para uma estratégia composicional eficaz no que concerne à interação, é na verdade através do contraste que as identidades das transferências espectrais em composições mistas podem ser avaliadas pelo ouvinte (MENEZES 2002 p.305).

Não resta dúvida de que um equilíbrio entre passagens baseadas em contraste e outras em semelhanças confere variedade e dinamismo ao discurso da interação. Entretanto uma pergunta mais fundamental não é aventada pela proposição de Menezes: qual a implicação de se usar contraste em vez de similaridade de conteúdos espectro-morfológicos dos materiais?

Uma peça elaborada perpetuamente por similaridades poderia não reter nossa atenção por muito tempo. Contrastes introduzem elementos dramáticos que podem se tornar essenciais para o desenvolvimento da narrativa. Porém qualquer drama requer uma resolução. Em outras palavras, um contraste introduz um problema que demanda uma solução. O contraste entre materiais instrumentais e eletrônicos nas obras eletroacústicas mistas pode ser um recurso fundamental para a criação da forma da obra, assim como o contraste entre tonalidades o foi para a forma da Sonata clássica, ou o contraste entre solo e ripieno para a forma do concerto barroco.

A composição da peça do Exemplo 5 respondeu a um desafio do musicólogo Hanns-Bertold Dietz. Ele propôs o problema de escrever uma peça para piano e sons eletrônicos

que partisse das iniciais HBD do seu nome. Depois ouvir uma execução da peça, ele me relatou que ficara algo desconcertado pelos sons dos ruídos que lembravam buzinas e vidros quebrados que aparecem logo na seção de exposição temática. Ele não imaginava ser possível contrastar o tema que propusera a um material tão heterogêneo. Entretanto reconhecia que ao final da peça tudo parecia fazer sentido. Lembrei a ele que cada época histórica coloca o problema do contraste de modo diferente. Em nossa época, somente um contraste tão radical como a diferença de naturezas espectromorfológicas entre os materiais representaria um desafio que merecesse ser enfrentado.

Naquela peça a solução do problema do contraste se dá por similaridade dos materiais aproximações sucessivas na instrumentais e eletrônicos. Ao longo da peça aparecem passagens no registro agudo do piano que se relacionam com os sons de vidro transformados. As buzinas eletrônicas se deslocam para o grave para sugerir semelhanças com acordes percutidos na região grave do piano, e assim por diante. A fraseologia gestual, predominantemente paratática exposição, gradativamente dá lugar a frases organizadas sintaticamente como texturas. gestos figurações desenvolvidas. repetição de materiais inicialmente Α contrastantes, em outras situações que revelam relações de similaridade entre eles, confere sentido ao discurso.

## Problemas musicais decorrentes da tecnologia

Vimos acima que para superar o problema criado pelo contraste entre materiais é necessário encontrar relações de semelhança convincentes entre ele, sob pena de enfraquecimento da linguagem da obra. Entretanto esse não é o único perigo que pode surgir em nosso caminho. Smalley (1997, p.111) nos alerta para o problema da "escuta tecnológica:"

<u>Perigo Nº 1:</u> deixar que a audição da técnica de produção prevaleça sobre a audição das relações musicais.

Este é um problema que surge frequentemente em laboratório ao se criar muitas variações do mesmo material. No contexto da experimentação, as variantes podem parecer suficientemente contrastantes. Entretanto, combinadas numa composição, é comum que as diferenças se tornem irrelevantes. O problema pode estar em focar a percepção apenas nas diferenças entre as técnicas de produção. O compositor eletroacústico treina ouvido eliminar seu para referencialidade dos materiais e a prestar atenção nas estruturas internas. Acontece que, para a percepção regular, a morfologia genérica do material pode ser mais forte do que muitas das sutis variações espectrais que somos capazes de produzir em estúdio. Podemos não estar gerando coisas diferentes que parecem semelhantes, mas apenas cópias da mesma coisa.

Um outro problema correlato é confiar demais nas técnicas de efeitos para produzir variações de um material. É comum que, em vez de ouvirmos a transformação de um som, percebamos apenas o recurso técnico empregado transformação. Um caso corriqueiro é o emprego de retardos. Simplesmente injetando e reproduzindo um material numa linha de retardo, quase sempre apenas produzimos uma desagradável sensação de *dejá vu* pois não adicionamos nada ao objeto repetido. O que prevaleceu para o ouvido foi a própria técnica do retardo, não a variação do som. A solução clássica desse problema é nunca confiar num único processamento para se produzir uma variação de um material. O problema acima apontado poderia desaparecer se, ao mesmo tempo em que aplicamos o retardo, houvesse, por exemplo, uma filtragem espectral progressiva do material, a adição gradual de reverberação e uma modificação gradativa do perfil dinâmico.

Regra Prática Nº 2: Multi-processamentos ajudam a combater a escuta tecnológica

A escuta tecnológica ocorre, portanto, quando nos focamos apenas no método de produção do som, em vez de prestarmos atenção nos múltiplos sentidos do objeto sonoro. Outra ilusão da escuta tecnológica decorre da inexistência de

virtuosidade no som virtual. Máquinas podem tocar qualquer coisa, numa velocidade programada qualquer, independentemente da complexidade resultante. Humanos tem limites. Atingir ou superar esses limites tem um sentido quando uma pessoa executa uma passagem difícil. A mesma passagem tocada por uma máquina não cria sensação de virtuosidade. Entretanto é possível produzir a ilusão de extensão da virtuosidade humana se tornarmos difusas as arestas entre o som instrumental e o eletrônico. Obras clássicas desde a *Música su due Dimensioni* de Maderna, passando por *Kontakte* de Stockhausen e os diversos *Sincronismos* de Davidovsky contém inúmeros exemplos desse tipo de efeito mágico.

<u>Perigo №2:</u> O "reflexo automático" na produção de variações de sons captados ao vivo.

eletrônica sobre meios fixos (isto é, a que usa materiais pré-

uma espécie de máquina de respostas automáticas, um mecanismo newtoniano em que a toda ação corresponde uma reação previsível. Esse tipo de diálogo em que um fala e o outro sempre responde com uma variação do que o primeiro falou, logo incorre no problema da escuta tecnológica porque passamos a ouvir a técnica de produção do resultado e não

A discussão sobre as vantagens e as desvantagens da

gravados em tape, CD, ou arquivo digital, e que demanda do intérprete a habilidade de tocar ao vivo em sincronia com a parte fixa) versus eletrônica ao vivo, parece datada no atual estágio tecnológico, uma vez que a primeira pode ser vista como um caso trivial da segunda. Entretanto o *live eletronics* oferece um outro tipo de perigo de escuta tecnológica que ocorre quando repetimos automaticamente um processo do tipo, captar o som de um instrumento ao vivo, modificá-lo num computador e devolvê-lo alterado pelos auto-falantes. Criamos

mais o resultado em si.

## Referências

ANTOKOLETZ, Elliott. *Twentieth-century music.* Englewood Cliffs: Prentice-Hall, 1992.

EMMERSON, Simon. The relation of language to materials. In: \_\_\_\_\_\_. (org.), *The language of electroacoustic music*, p.17–39. Nova York: Harwood Academic, 1986.

FERNEYHOUGH, Brian. Le temps de la figure. In: *Entretemps*, v. 3, Paris: J. C. Lattès, 1987.

FERNEYHOUGH, Brian. Shattering the vessels of received wisdom: in conversation with James Boros (1990). In: *Collected writings*. Amsterdam: Harwood. 1995.

GILLESPIE, Brent. Haptics. In: COOK, Perry (org.), *Music, cognition and computerized sound: an introduction to psychoacoustics*, p.229–246. Cambridge, Mass: MIT Press, 1999.

HATTEN, Robert. *Musical meaning in Beethoven: markedness, correlation, and interpretation.* Bloomington: Indiana University Press. 1994.

KENDALL, Gary. Juxtaposition and non-motion: Varèse bridges early modernism to electroacoustic music. *Organised Sound*, v. 11, n. 2 (2006), p. 159–171.

MENEZES, Flo. For a morphology of interaction. *Organised Sound*, v. 7, n. 3 (2002), p. 305-311.

SANTAELLA, Lúcia. *Matrizes da linguagem e pensamento: sonora, visual e verbal.* São Paulo: Iluminuras. 2001.

SCHAEFFER, Pierre. Traité des objets musicaux. Paris: Seuil. 1966.

SMALLEY, Denis. Spectro-morphology and structuring processes. In: EMMERSON, Simon (org.), *The language of electroacoustic music*, p. 61–93. Nova York: Harwood Academic. 1986.

SMALLEY, Denis. Spectromorphology: explaining sound-shapes. *Organised Sound*, v. 2, n. 2 (1997), p. 107-129.

THORESON, Lasse; HEDMAN, Andreas. Spectromorphological analysis of sound objects: an adaptation of Pierre Schaeffer's typomorphology." *Organised Sound*, v. 12 (2007), p. 129-141.

XAVIER, Ismail. *O discurso cinematográfico: a opacidade e a transparência*. 3 ed. São Paulo: Paz e Terra. 2005.

XENAKIS, Iannis. *Formalized music*. Hillsdale, NY: Pendragon Press. 1992.

#### Performance musical em rede

ÁLVARO BARBOSA

o âmbito da performance musical mediada por tecnologia informática (sobretudo em redes de comunicação), é bem conhecido que a latência tem um enorme efeito disruptivo na prática musical colaborativa, tanto no nível da resposta musical do próprio performer como na resposta musical global resultante da contribuição de outros performers num ato colaborativo.

É também extensamente reconhecido que o advento da internet e a possibilidade de comunicações acústicas globais a longa distância trouxeram perspectivas de concretização de performances musicais geograficamente deslocalizadas, acessíveis a uma comunidade mundial. No entanto, neste cenário, a latência de comunicação é ainda maior que a latência introduzida pela propagação sonora em grandes espaços ou aquela que é causada por uma computação intensa.

Numa entrevista em vídeo no ensaio *Soft Music*, de Golo Föllmer, Atau Tanaka expressa uma visão inspiradora sobre este tópico:

Considero a latência na internet bastante interessante e penso nisso como uma espécie de acústica única deste media [...] mais do que tocar música já existente nesta nova base temporal, o que me parece interessante é tentar encontrar uma linguagem musical que funcione nesta linha temporal [...] se um som demora meio segundo para ir de Paris a Nova York e outro meio segundo para voltar, então podemos criar uma música adaptada a esta acústica. (FÖLLMER, 2001)

A proposta de que a latência na internet é uma característica acústica inerente a este meio e que os compositores podem criar música tendo em conta este fato vai de encontro à noção recorrente de que a adaptação da música ao meio onde é tocada leva ao aparecimento de novidades estilísticas. Um exemplo notório deste conceito é a música policoral veneziana do final da Renascença e princípio do Barroco, originada pela acústica peculiar do espaço arquitetônico da Basílica de São Marcos, em Veneza, causada pela latência de propagação sonora no espaço físico em função da distância entre as galerias opostas dos coros. (REESE, 1954) (BUKOFZER, 1947).

Ainda assim, a latência é essencialmente vista como um elemento disruptivo especialmente nas formas musicais mais tradicionais, que são na sua maioria conduzidas pelo ritmo e pela melodia e por conseguinte requerem uma sincronização perfeita de modo a que o performer obtenha uma consciência da acústica musical em tempo real.

Seja ultrapassando os seus efeitos desconcertantes, ou adaptando-a enquanto elemento construtivo nas práticas musicais, a latência implica diferentes abordagens em relação à *performance* musical, especialmente no nível do gesto humano e do seu efeito direto no desenvolvimento do resultado musical.

Neste capítulo aborda-se a presença da latência extrema na performance musical fundamentada na pesquisa relacionada com o uso da internet numa perspectiva de colaboração musical com base em experiências laboratoriais levadas a cabo no MTG-Barcelona e no CITAR-Porto desde 2003, que analisam de que forma a pratica musical é afetada pela latência de comunicação e o que pode ser feito para diminuir este efeito desconcertante.

## A percepção da acústica da internet

O grupo de pesquisa *SoundWire* no Centro para a Pesquisa Computacional em Música e Acústica (CCRMA - Center for Computer Research in Music and Acoustics) na Universidade de Stanford, dirigido por Chris Chafe, publicou nos últimos anos vários artigos de pesquisa tratando das implicações das condições da rede na comunicação acústica (CHAFE, 2000, p. 159) (CHAFE e LEISTIKOW, 2001).

182

Em particular no artigo "Physical Model Synthesis with Application to the Internet Acoustics" (CHAFE, WILSON e WAILLING, 2002) Chafe descreve como os modelos físicos de instrumentos musicais para sistemas distribuídos têm sido usados para produzir *pings* acústicos em ligações à internet entre dois servidores de rede,¹ partindo da observação de que quando as ondas sonoras se propagam através da acústica da internet elas se comportam da mesma maneira que no ar, água ou ao longo duma corda tensa.

A ideia de "ouvir o som de uma rede" é uma visão estimulante de como a latência da rede pode ser vista como a principal propriedade para a caracterização da acústica da internet.

Em colaboração com o artista Greg Niemeyer, esta mesma ideia levou à instalação sonora experimental no MOMA de São Francisco, intitulada *Ping* (CHAFE e NIEMEYER, 2001), em que a síntese por modelo físico é utilizada para a sonificação dos dados da Internet.

A Internet apresenta diferentes características que podem afetar qualquer processo colaborativo. O sincronismo em tempo real é inquestionavelmente central na prática musical, e de forma geral a latência é o principal impedimento para a colaboração musical em tempo real.

Este problema está presente em muitos outros contextos além do da comunicação em redes de longa distância, como, por exemplo, nas placas de som do computador ou nos sistemas de amplificação sonora nos grandes auditórios, onde o som dos monitores de áudio traseiros tem de sofrer atraso de propagação de sinal elétrico para poder corresponder à fase do som que vem do palco e que sofreu um atraso maior durante a sua propagação pela atmosfera.

<sup>&</sup>lt;sup>1</sup> Nesta experiência são utilizados modelos físicos de síntese de instrumentos de cordas que são excitados pelo sinal resultante da execução de um comando *ping* que produz uma oscilação com um período correspondente à latência de rede e determinando a variável correspondente ao comprimento da corda do modelo de acordo com o respectivo comprimento de onda deste sinal.

Um exemplo usualmente apresentado como ilustrativo do efeito disruptivo da latência acústica provocado pela propagação do som na atmosfera é o cenário onde dois músicos tentam tocar juntos, cada um em lados opostos de um estádio de futebol (cerca de 120 metros de distância entre eles). O som demorará cerca de 35 ms (considerando o nível do mar, a temperatura de 15ºC e a velocidade do som a 340 m/s) para ir de um músico ao outro, e a resposta musical individual (ida e volta), para cada músico perceber a reação do outro será o dobro desse valor (70 ms). Esses valores são muito altos e seria muito complicado chegar a uma performance equilibrada e sincronizada nestas condições.

## A tolerância à latência na performance musical

Para que o ouvido humano percepcione dois sons como simultâneos, eles não podem estar separados temporalmente em mais do que 20 ms (HIRSH, 1959, p. 759), o que significa que para uma percepção sonora simultânea numa performance bilateral este limiar deve ser por volta de 40 ms (o período de tempo que demora para que um performer perceba a reação do outro à sua ação).

Note-se que a percepção de dois sons diferentes tocados simultaneamente depende fortemente das características do som (timbre, altura ou intensidade), estilo musical e outros tipos de resposta, tais como os estímulos visuais ou físicos. Ainda assim, um limiar de 20 ms é suficientemente grande para caracterizar o pior caso.

De fato, uma série de experiências foram levadas a cabo com o propósito de determinar quanta latência de comunicação poderia ser tolerada entre músicos de forma a manter uma performance síncrona.

Os significantes resultados da pesquisa conduzida em 2002 na Universidade de Stanford por Nathan Shuett (SHUETT, 2002) estabeleceram experimentalmente um limiar para a performance em conjunto (EPT - Ensemble Performance Threshold) para música de estrutura rítmica com pulsações situadas entre os 20-30ms, que estão de acordo com os

resultados da pesquisa feita por Nelson Lago em 2004 (LAGO e KON, 2004, p. 33) na Universidade de São Paulo.

No contexto da transmissão de áudio através de rede de computadores, considerando os avanços na performance de banda larga e compressão de informação, podemos ser levados a pensar que a latência da rede é uma imposição tecnológica que pode ser ultrapassada num futuro próximo, e por isso poderia ser desnecessário estudar formas de diminuir os seus efeitos perturbadores na performance musical tradicional.

Ainda que não consideremos a latência extrema introduzida na comunicação por satélite, ou que a emergente tecnologia móvel tem taxas de transmissão de dados muito baixas, pode ser demonstrado pelas leis da física que em nível global existem limites que irão sempre implicar em níveis de latência mais altos do que o limiar mínimo tolerado em colaborações musicais em tempo real.

184

Tomando como exemplo uma conexão *peer-to-peer* abstrata por via do percurso mais curto possível entre dois pontos opostos no planeta, digamos, Santiago do Chile e Moscou, temos uma distância aproximada de 14.141 km. Mesmo com uma idealizada transferência de dados por fibra óptica, sem perdas, à velocidade da luz (299.792,458 km/s) e largura de banda ilimitada, a latência bidirecional seria aproximadamente 94,3 ms, o que é muito maior que o limiar mínimo tolerável.

Além disso, a latência tem uma natureza altamente variável e imprevisível, criando erros de base temporal, dessequenciando e ainda provocando a perda parcial do conteúdo, resultando em sérios constrangimentos para o controle da performance. Ainda assim, um enorme esforço está a ser feito na comunidade científica para diminuir estas limitações, através do aumento da largura de banda, da capacidade de compressão de dados e pelo uso de técnicas de transmissão baseadas no conteúdo.



**Fig. 1:** Cenário de uma comunicação idealizada entre duas cidades geograficamente opostas em temos globais.

No entanto, para o caso de redes de curta distância ou ainda redes de longa distância em territórios geograficamente demarcados (um país ou mesmo um continente) pode ser esperado que num futuro próximo a latência de rede venha a ser reduzida para valores que não representarão um impedimento para comunicações acústicas em tempo real pela internet

## Tempo e dinâmica adaptativos em função da latência

Algumas das pesquisas referentes aos efeitos da latência na precisão temporal na performance colaborativa vão além de estabelecer um EPT para um cenário geral de sincronização rítmica.

Um trabalho publicado em 2004 por Chris Chafe e Michael Gurevish (CHAFE e GUREVISH, 2004), resultante duma experiência conduzida no CCRMA, mostra que pares de sujeitos que tentam sincronizar um ritmo constante batendo palmas tendem a diminuir o ritmo com o aumento da latência.

Do mesmo modo, uma experiência levada a cabo pelo Junho de 2004 no Departamento de Som e Imagem da Universidade Católica Portuguesa (UCP) destinou-se, entre outras metas, a estudar a relação entre tempo (ritmo) e latência.

Na experiência, foram aplicadas condições de latência de rede simuladas a uma performance de quatro músicos diferentes a tocar *standards* do repertório de jazz com quatro instrumentos diferentes (baixo, percussão, piano e guitarra).





**Fig. 2:** Experiência de tolerância à latência num ambiente de estúdio simulado conduzida por Álvaro Barbosa e Alexander Carôt.

A primeira parte desta experiência consistiu em determinar o máximo individual de tolerância à latência aplicada à resposta musical individual do instrumento de cada músico.

Com esta finalidade, foi montado um sistema em estúdio para que os músicos ouvissem a resposta dos seus próprios instrumentos através de fones de ouvido com latências variáveis.

As performances eram sincronizadas com um metrônomo durante várias passagens com andamentos diferentes (BPMs - *beats per minute*). Para cada passagem a latência da resposta musical era aumentada até que o músico não fosse capaz de manter uma performance síncrona.

O gráfico e a tabela seguintes (ver Gráf. 1, Tab. 1) mostram os resultados dessa experiência preliminar.

Da análise dos resultados fica claro que, independentemente das capacidades instrumentais de cada músico, todos foram capazes de tolerar maior *feedback* para tempos mais lentos, conforme se pode confirmar no gráfico seguinte.

A única exceção para a tendência decrescente destas curvas ocorre quando o percussionista atinge os 160 BPMs, o que está relacionado com uma sobreposição síncrona sobre a estrutura rítmica da música, em conjunto com o fato de que para instrumentos de percussão é muito difícil isolar totalmente o performer do som direto do instrumento. Daqui se depreende uma relação inversa entre tempo musical e tolerância à latência.

Para uma validação adicional a esta hipótese foi realizado um teste com utilizadores no contexto desta experiência (ver Fig. 3). Um questionário *online* foi submetido a 32 sujeitos com um perfil dominante de estudantes de música da Escola de Artes da Universidade Católica Portuguesa (53% com formação em música; 28% conseguem tocar um instrumento musical; 19% não têm qualquer formação musical).

O questionário consistia em classificar a precisão de uma performance musical do bem conhecido *standard* de jazz "Sunny", composição de Bobby Hebb, interpretado por diferentes pares de instrumentos (baixo/percussão; baixo/guitarra; baixo/piano) e para diferentes combinações de tempo para uma latência de comunicação fixa entre músicos de 30 ms (35ms no caso do dueto baixo/piano).

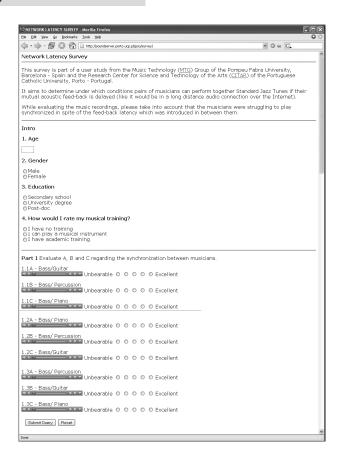


Fig. 3: Questionário online para avaliar a relação entre tempo musical e latência de comunicação.

O tema musical "Sunny" foi escolhido uma vez que foi uma das músicas tocadas na anterior experiência *Stanford-McGill* (13 de Junho de 2002), em que ficou claro, a partir da observação empírica da documentação em vídeo, que, em alguns momentos, os músicos não conseguiam manter uma performance bem sincronizada. Desta forma, o material gravado nas sessões do Porto poderia ser comparado com o da experiência *Stanford-McGill*.

189

	Músicos								
	Baixo	Percussão	Percussão Guitarra						
80	-	85 ms	180 ms	-					
100	250 ms	75 ms	130 ms	165 ms					
110	-	-	-	160 ms					
120	-	70 ms	-	150 ms					
130	225 ms	-	100 ms	150 ms					
140	-	60 ms	-	130 ms					
150	150 ms	-	60 ms	-					
160	-	65 ms	-	-					
170	125 ms	-	-	-					
190	100 ms	-	-	-					

Tempo (BPMs)

**Tab. 1:** Tolerância máxima à latência para cada músico tocando a andamentos diferentes (BPMs *Beats Per Minute*).

Os resultados presentes na tabela seguinte mostram que os sujeitos consideraram que as performances com aproximadamente a mesma latência eram geralmente melhores para tempos musicais mais baixos (100 BPM), independentemente dos instrumentos e dos performers (ver Tab. 2).

Date	IP Address	Age	Gender	Education	Musical Training	1.1A	1.1B	1.1C	1.2A	1.2B	1.2C	1.3A	1.3B	1.3C
27-04-2005 11:4	0 172.20.80.60	20	M	University degree	Academic training	3	0	2	1	0	1	0	2	1
20-04-2005 14:5		25	F	University degree	Academic training	3	0	2	0	1	3	0	2	0
19-04-2005 12:0	7 172.20.80.60	38	F	University degree	Academic training	2	1	3	0	0	2	0	1	2
29-06-2005 16:1	9 172.20.80.60	31	M	University degree	Academic training	0	0	3	0	0	1	2	0	1
18-04-2005 16:2		23	M	University degree	Academic training	4	1	3	0	0	2	2	2	3
18-04-2005 16:2	0 172.20.80.60	26	M	University degree	Academic training	3	1	2	1	0	3	1	0	3
02-05-2005 16:2	5 172.20.80.60	25	F	University degree	Academic training	3	1	1	1	2	1	2	2	0
02-05-2005 16:5	8 172.20.80.60	25	M	University degree	Academic training	2	1	0	1	0	1	1	0	1
03-05-2005 15:0	0 172.20.80.60	23	F	University degree	Academic training	2	3	3	2	1	2	0	2	1
03-05-2005 17:2	6 172.20.80.60	32	M	University degree	Academic training	2	1	2	3	1	2	2	2	1
06-05-2005 12:3	7 172.20.80.60	28	M	University degree	Academic training	1	0	0	1	0	2	2	0	1
07-05-2005 11:5	7 84.143.179.74	45	M	Post-doc	Academic training	1	2	0	1	1	2	1	0	1
11-05-2005 19:1	7 193.145.55.204	28	M	University degree	Academic training	3	1	0	2	1	0	1	1	0
11-05-2005 20:3	2 193.145.55.204	29	M	University degree	Academic training	2	1	3	1	2	2	3	2	2
12-05-2005 14:5	9 172.20.80.60	24	F	University degree	Academic training	2	0	2	1	1	0	1	2	0
19-05-2005 10:4	5 172.20.80.60	28	M	University degree	Academic training	2	0	1	0	1	3	1	1	0
24-05-2005 11:1	8 172.20.80.60	24	F	University degree	Academic training	3	2	2	1	0	2	1	2	2
					Average:	2.24	0.88	1.71	0.94	0.65	1.71	1.18	1.24	1.12
													-	
28-04-2005 10:0	9 172.20.80.60	23	M	University degree	Can play musical instrument	3	1	3	0	2	2	3	2	2
28-04-2005 10:3	4 172.20.80.60	26	М	University degree	Can play musical instrument	3	2	3	1	2	3	2	1	2
28-04-2005 11:5	5 172.20.80.60	34	M	University degree	Can play musical instrument	3	1	3	0	1	2	1	2	2
03-05-2005 10:1	4 194.117.24.10	31	M	University degree	Can play musical instrument	3	0	1	3	1	2	2	2	1
03-05-2005 13:0	7 172.20.80.60	36	F	University degree	Can play musical instrument	3	3	1	1	0	2	1	1	3
07-05-2005 17:4	6 172.20.80.60	24	M	Secondary School	Can play musical instrument	3	2	2	3	1	1	0	3	1
11-05-2005 18:0	7 193.145.55.204	37	M	University degree	Can play musical instrument	2	1	2	1	1	1	1	0	1
11-05-2005 20:0	3 193.145.55.204	25	M	University degree	Can play musical instrument	1	2	1	1	1	1	1	2	1
16-05-2005 15:1	5 141.83.78.62	28	F	University degree	Can play musical instrument	4	1	2	2	1	2	3	3	1
					Average:	2.79	1.44	2	1 33	1 11	1 79	1.56	1 79	1.56
					Average	2,70	1,44		1,00	1,111	1,70	1,50	1,70	1,50
23-04-2005 16:0	8 172.20.80.60	33	М	University degree	No Training	3	3	2	4	1	1	1	2	0
30-04-2005 17:5	4 192.35.246.5	36	М	University degree	No Training	1	2	0	0	0	2	1	1	3
02-05-2005 11:1		34	F	University degree	No Training	3	3	2	1	3	0	0	4	4
	3 193.145.56.194	26	F	University degree	No Training	3	3	1	2	1	4	2	2	2
•				, , , , ,	Average	2,5	2.75	1.25	1.75	1.25	1.75	1	2.25	2.25
					Average.	2,3	2,/3	1,23	1,73	1,20	1,73		2,20	2,23
					Final Average:	2.5	1 69	1 65	1 34	1	1 74	1.24	1 75	1 64

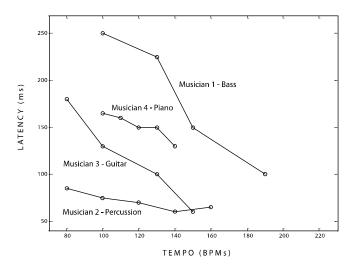
190

**Tab. 2:** Resultados da avaliação da relação tempo musical/latência da comunicação agrupada por nível de formação musical.

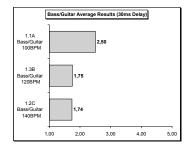
Os gráficos seguintes mostram de forma mais clara a dependência direta entre o tempo musical e a tolerância ao efeito disruptivo da latência, o que confirma a hipótese inicial (ver Gráf. 1 e Gráf. 2).

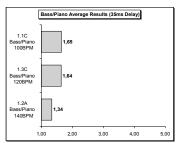
A dependência direta entre o tempo musical e a tolerância ao efeito disruptivo da latência, neste caso específico da colaboração musical (performance standard de jazz), pode ser vista como um conceito mais geral de Adaptação Temporal à Latência (LAT - Latency Adaptive Tempo).

O princípio base de aplicação da LAT consiste numa função para sistemas de comunicação acústica de rede, cujo tempo musical (tipicamente dirigido pelo som de um metrônomo) se adapta, transpondo-se para o valor máximo tolerado pela menor "tolerância à latência" dum músico participante na sessão performativa. Esta adaptação dinâmica é baseada em medidas da latência em tempo real entre pares.

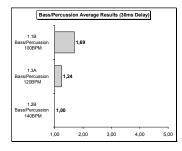


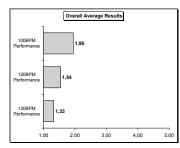
**Gráf. 1:** Auto-teste de tolerância à latência para cada performance.





**Gráf. 2:** Resultados da avaliação da relação tempo musical/latência da comunicação no caso dos duetos baixo/guitarra e baixo/piano.





**Gráf. 3:** Resultados da avaliação da relação tempo musical/latência da comunicação no caso dos duetos baixo/percussão e a média final

As variáveis de entrada desta função são os perfis de desempenho dos músicos e o valor de latência num dado momento. A resposta da função LAT será o valor do tempo (tipicamente em BPMs) que é menos disruptivo para a prática do grupo musical. A LAT permite aos músicos ensaiar música tão depressa (em termos de tempo musical) quanto a velocidade de conexão da sua rede lhes permite.

A ideia genérica de um instrumento musical em rede que se adapta dinamicamente à latência da conexão de internet foi implementada originalmente por Jörg Stelken no software peerSynth (STELKENS, 2003). PeerSynth é um sintetizador peerto-peer que suporta múltiplos utilizadores remotos na internet, medindo a latência entre cada conexão ativa e reduzindo dinamicamente o volume do som da contribuição de cada utilizador, na paisagem sonora que está a ser criada, proporcionalmente ao valor da latência medido na sua conexão. Stelkens seguiu uma metáfora do mundo real, onde, de fato, a intensidade de um som de uma fonte sonora diminui com o aumento da distância do receptor, o que também implica um aumento da latência de comunicação acústica. Uma abordagem similar da adaptação dinâmica à latência (LAD - Latency Adaptive Dynamics) foi seguida no sistema AALIVENET (SPICER, 2004).

## Resposta individual com latência

Outro resultado obtido a partir das experiências com a simulação de latência de comunicação acústica na Escola das Artes da Universidade Católica Portuguesa foi uma tipologia de resposta musical que realça a tolerância individual à latência.

Foi empiricamente observado que, com a prática, os músicos tendem a melhorar as suas capacidades para tocar o seu instrumento musical quando a sua resposta acústica individual sofre uma latência. Esta ideia é reforçada pelos resultados apresentados no Gráfico 1, no qual podemos observar diferentes níveis de tolerância para a resposta musical individual para músicos com diferentes capacidades instrumentais.

Isto leva também à percepção de que uma melhor tolerância à latência é alcançada se, em vez de formarmos grupos em que cada músico recebe uma resposta acústica direta do seu instrumento misturado com a resposta musical global com latência, cada músico ouvir a sua resposta musical individual atrasada temporalmente, mas em conjunto e sincronizada com os outros músicos. Este conceito é definido como resposta individual com latência (IDF - Individual Delayed Feedback).

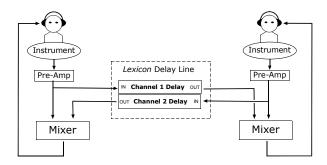


Fig. 4: Topologia da resposta musical individual sem latência.

Fig. 5: Topologia da resposta musical individual com latência.

As figuras anteriores ilustram a montagem em estúdio usada para as sessões de gravação entre pares de músicos. A mesma canção foi gravada com o mesmo tempo e latência, mas usando a topologia de resposta musical individual sem latência numa sessão e uma topologia de resposta musical individual com latência noutra sessão.

A canção gravada com estas duas topologias de resposta musical foi a bem conhecido *standard* de jazz "Cantaloupe Island", de Herbie Hancock, com um andamento de 120 BPM para uma latência de comunicação de 35 ms.

Quatro pares diferentes de performances instrumentais foram gravadas: baixo / guitarra; baixo / percussão; baixo / piano; piano / percussão. Essas gravações foram usadas no seguinte questionário *online* com utilizadores:

**Fig. 6:** Questionário *online* para a avaliação das performances com resposta individual com latência.

Mais uma vez, o teste foi submetido a 32 sujeitos com um perfil dominante de estudantes de música da Escola das Artes da Universidade Católica Portuguesa (53% com formação em música; 28% conseguem tocar um instrumento musical; 19% sem nenhuma formação musical). A canção A corresponde sempre à topologia de resposta musical individual sem latência, e a canção B corresponde à topologia de resposta musical individual com latência. Os resultados são apresentados na tabela seguinte:

Date	IP Address	Age	Gender	Education	Musical Training	2.1	2.2	2.3	2.4
27-04-2005 11:40	172.20.80.60	20	M	University degree	Academic training	В	В	В	В
20-04-2005 14:58	172.20.80.60	25	F	University degree	Academic training	В	В	В	В
19-04-2005 12:07	172.20.80.60	38	F	University degree	Academic training	В	В	В	В
29-06-2005 16:19	172.20.80.60	31	M	University degree	Academic training	В	В	В	В
18-04-2005 16:24	172.20.80.60	23	M	University degree	Academic training	В	В	В	В
18-04-2005 16:20		26	М	University degree	Academic training	В	В	В	В
02-05-2005 16:25	172.20.80.60	25	F	University degree	Academic training	В	В	В	В
02-05-2005 16:58	172.20.80.60	25	М	University degree	Academic training	В	В	В	В
03-05-2005 15:00	172.20.80.60	23	F	University degree	Academic training	В	В	В	В
03-05-2005 17:26	172.20.80.60	32	М	University degree	Academic training	В	В	Α	Α
06-05-2005 12:37	172.20.80.60	28	М	University degree	Academic training	В	В	В	В
07-05-2005 11:57	84.143.179.74	45	M	Post-doc	Academic training	В	В	В	В
11-05-2005 19:17	193.145.55.204	28	M	University degree	Academic training	В	В	В	В
11-05-2005 20:32	193.145.55.204	29	M	University degree	Academic training	В	В	В	В
12-05-2005 14:59	172.20.80.60	24	F	University degree	Academic training	В	В	В	В
19-05-2005 10:45	172.20.80.60	28	М	University degree	Academic training	В	В	В	Α
24-05-2005 11:18	172.20.80.60	24	F	University degree	Academic training	В	В	В	В
28-04-2005 10:09	172.20.80.60	23	M	University degree	Can play musical instrument	В	В	Α	В
28-04-2005 10:34	172.20.80.60	26	M	University degree	Can play musical instrument	В	В	В	Α
28-04-2005 11:55	172.20.80.60	34	M	University degree	Can play musical instrument	В	В	В	В
03-05-2005 10:14	194.117.24.10	31	M	University degree	Can play musical instrument	В	В	В	В
03-05-2005 13:07	172.20.80.60	36	F	University degree	Can play musical instrument	В	В	В	В
07-05-2005 17:46	172.20.80.60	24	M	Secondary School	Can play musical instrument	В	В	Α	В
11-05-2005 18:07	193.145.55.204	37	М	University degree	Can play musical instrument	В	В	В	В
11-05-2005 20:03	193.145.55.204	25	M	University degree	Can play musical instrument	В	В	В	В
16-05-2005 15:15	141.83.78.62	28	F	University degree	Can play musical instrument	В	В	В	В
23-04-2005 16:08	172.20.80.60	33	М	University degree	No Training	В	В	В	В
30-04-2005 17:54	192.35.246.5	36	М	University degree	No Training	Α	В	Α	Α
02-05-2005 11:19	172.20.80.60	34	F	University degree	No Training	В	В	В	В
10-05-2005 11:23		26	F	University degree	No Training	В	В	В	В
					A:	1	0	4	4
					B:	31	31	28	28
					5.				

**Tab. 3:** Resultados do questionário online na avaliação da Resposta Individual com Latência

Em qualquer dos casos mais de 85% dos sujeitos considera que a topologia IDF (Canção B) produz resultados melhores.

Baseado nesta corroboração da suposição de que uma topologia IDF permite melhores tolerâncias individuas à latência, o co-autor destas experiências, Alexander Carôt, implementou uma ferramenta para resposta musical com latência na sua aplicação para comunicação acústica de baixalatência na internet, intitulada *Soundjack* (CARÔT, 2004).

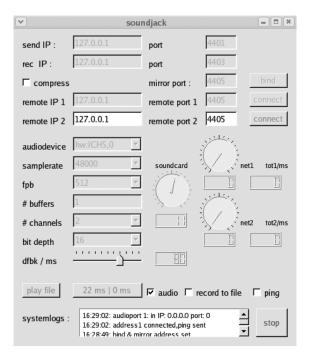


Fig. 7: Interface do Soundjack de Alexander Carôt

A interface permite aos utilizadores fazerem corresponder manualmente o valor da sua resposta individual com latência à latência da sessão, movendo o *slider* "dfbk/ms".

Os conceitos do IDF e do LAT são igualmente aplicados no sistema de objetos sonoros partilhados (BARBOSA, 2005), um ambiente colaborativo desenhado para fornecer uma interface gráfica orientada para a manipulação individual do som que funciona como um cliente de um espaço acústico partilhado na internet, em que a resposta acústica e visual, entre outras, é sujeita a latências extremas, variáveis com as condições de velocidade da rede.

**Fig. 8:** Software concebido por Álvaro Barbosa: *Objetos Sonoros Partilhados* (PSOs - *Public Sound Objects*)

## Conclusões

Quando enquadrada numa concepção tradicional da música, a performance colaborativa requer comunicação em tempo real entre performers. Contudo, a latência de comunicação tem um efeito disruptivo na sincronização musical. Neste caso os músicos não têm uma resposta imediata aos seus gestos performativos. Neste capítulo foi apresentada uma análise de questões perceptivas relacionadas com a latência extrema.

Os conceitos de adaptação temporal e dinâmica à latência (LAT e LAD) derivam de uma experiência e avaliação que demonstrou uma dependência direta entre tempo musical e tolerância ao efeito disruptivo da latência num caso específico de colaboração musical (performance *standard* de jazz). Este conceito foi implementado pelo autor no projeto *Objetos Sonoros Partilhados* fornecendo uma melhoria significativa na percepção da performance individual em condições extremas de latência na resposta acústica.

Para além disso foi introduzido o conceito de resposta individual com latência (IDF), baseado na demonstração experimental que melhor tolerância à latência é alcançada quando um musico recebe a resposta acústica do seu instrumento integrada com a resposta com latência dos outros performers, em vez da resposta direta individual do seu próprio instrumento. Neste caso cada músico percepciona a sua própria resposta individual com latência, juntamente e

sincronizada com a resposta dos outros músicos. O IDF foi utilizado no aplicativo *Soundjack*, para comunicação acústica de baixa-latência pela internet, e no sistema *Objetos Sonoros Partilhados*.

As noções de LAT e LAD assim como a IDF são conceitos incorporáveis no contexto da performance musical *online*. Com base nestes princípios, os músicos saberão que a fim de improvisar *online*, têm de melhorar as suas capacidades tocando com uma resposta acústica com latência e que serão aptos a tocar quanto mais rápido quanto a sua conexão de internet lhes permitir.

Com o emergir de aplicações musicais que incorporam a latência da internet como uma parte funcional do sistema, em vez de tentar cancelá-la, possivelmente chegaremos a estilos musicais com menos estruturas rítmicas e com ataques e decaimentos mais lentos (ver capítulo 6 deste livro).

Adicionalmente, a tecnologia digital e o acesso a novos tipos de sensores permitem uma abordagem às interfaces de instrumentos musicais que pode ser concebida como raiz para mapear em tempo real os gestos humanos em parâmetros sonoros. No entanto, os mecanismos algorítmicos de geração de som permitem criar música em que o performer improvisador tenha uma maior consciência do comportamento musical de estruturas rítmicas e melódicas em contraponto com um modelo de performance musical nota por nota como resultado direto dos seus gestos (interfaces comportamentais). Neste sentido, os instrumentos digitais não só permitem uma interação orientada comportamento, como ao também aumentam a tolerância à latência.

### Referências

BARBOSA, Álvaro; CARDOSO, Jorge; GEIGER, Gunter. Network latency adaptive tempo in the public sound objects system. In: *Proceedings the International Conference on New Interfaces for Musical Expression*, Vancouver, 2005.

BUKOFZER, Manfred. *Music in the Baroque era*. Nova York: Norton, 1947.

CARÔT, Alexander. *Live music on the internet*. Lübeck: IT & Design, 2004

CHAFE, Chris; GUREVICH, Michael; GRACE, Leslie; SEAN, Tyan. Effect of time delay on ensemble accuracy. In: *Proceedings of the International Symposium on Musical Acoustics*, Nara, 2004

CHAFE, Chris; NIEMEYER, Greg. *Ping music installation - 2001*. Walker Art Center and San Francisco Museum of Modern Art: http://www.ccrma.stanford.edu/~cc/sfmoma/topLevel.html [consultado em 30 de Janeiro de 2010]

CHAFE, Chris; WILSON, Scott Robert; LEISTIKOW, Randal; CHISHOLM, Dave; SCAVONE, Gary. Simplified approach to high quality music and sound over IP. In: *Proceedings of the Digital Audio Effects Conference*, Verona, 2000.

CHAFE, Chris; WILSON, Scott Robert; WALLING, Daniel. Physical model synthesis with application to internet acoustics. In: *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, Orlando, 2002. IEEE - Signal Processing Society.

HIRSH, Ira. Auditory perception of temporal order. *Journal of the Acoustical Society of America*, v. 31, n. 6 (1959), p. 759.

LAGO, Nelson; KON, Fabio. The quest for low latency. In: *Proceedings of the International Computer Music Conference*, Miami, 2004

REESE, Gustave. Music in the Renaissance. New York: Norton, 1954.

SCHUETT, Nathan. *The effects of latency on ensemble performance*. Stanford, 2002. Monografia de Graduação – Stanford University.

SPICER, Michael. AALIVENET: An agent based distributed interactive composition environment. In: *Proceedings of the International Computer Music Conference*, Miami, 2004.

STELKENS, Jörg. PeerSynth: A P2P Multi-User Software with new techniques for integrating latency in real time collaboration. In: *Proceedings of the International Computer Music Conference*, Singapore, 2003.

#### Glossário

## preparado por JOSÉ FORNARI, DANIEL BARREIRO

**amplitude**: este termo pode se referir a dois aspectos de uma onda sonora: o seu valor instantâneo e o seu valor máximo (de pico ou em media). No primeiro caso, é equivalente ao valor do sinal de áudio como função do tempo; no segundo é uma medida proporcional à quantidade de energia em um sinal.

**amostra**: em áudio digital, amostra é o valor obtido de um sinal analógico no processo de amostragem. Uma sequência de amostras resulta em um sinal digital.

**amostragem**: é parte do processo de conversão analógica-digital de sinais de áudio. Nele, o sinal analógico é medido em intervalos regulares de tempo (período de amostragem), produzindo uma sequência de amostras de áudio. Por exemplo, o sinal digital em um CD de áudio é amostrado 44.100 vezes por segundo (ver *taxa de amostragem*).

aditiva (síntese): técnica de síntese baseada em uma soma de sinais senoidais de frequencias, amplitudes (e fases) distintas, produzindo uma onda sonora complexa. Geralmente este método é implementado com um grupo de osciladores em paralelo, cada um modelando um componente do som a ser gerado.

audição (limites): em frequência (20-20KHz); em intensidade (0 - 120 dB SPL)

analógico: um sinal contínuo no tempo e em amplitude.

ADC (analog-digital-conversion): conversão digital-analógica.

array: vetor (uma dimensão); matriz (duas ou mais dimensões).

bi-audição: localização espacial, azimute.

**Csound**: sistema de síntese e linguagem desenvolvido inicialmente por Barry Vercoe no MIT.

**DSP** (digital signal processing): processamento digital de sinais

digital: um sinal discreto no tempo e em amplitude.

DAC: (digital-analog convesion): conversão digital-analógica

**DFT** (Discrete Fourier Transform): Transformada discreta de Fourier

**EPT** (Ensemble Performance Threshold): limite de tolerância à latência de rede em performance.

**fase**: em termos de ondas senoidais a fase se refere à sua posição angular em um certo momento (o ângulo, ou argumento da função senoidal); o termo

também se refere ao desvio de fase, ou seja, à diferença de fase entre dois sinais de mesma frequência fundamental (e do mesmo tipo de onda), que tambem é relacionado com o retardo de sinais.

**filtro** (filtragem): processadores que em geral afetam a amplitude e/ou a fase de um sinal em diferentes frequências, cortando, amplificando, atenuando ou retardando sinais, dependendo do tipo utilizado.

low-pass filter: filtro passa-baixa high-pass filter: filtro passa-alta band-pass filter: filtro passa-banda

**frequência**: a medida de uma quantidade, fenômeno ou evento que se repete no tempo. No caso de ondas senoidais, é o inverso do período de tempo em que um ciclo se completa, ou o número de ciclos completos um dado período de tempo. Quando nos referimos a ondas periódicas complexas, geralmente usamos o termo frequência fundamental, que é mais específico para estes sinais. A unidade mais comum usada para frequência é o Hertz (Hz), que equivale a um ciclo por segundo.

FM synthesis: síntese por modulação de frequência.

ganho: um multiplicador aplicado ao sinal, amplificando ou atenuando este.

**IDF** (Individual Delayed Feed-Back): resposta individual atrasada com a performance coletiva.

ITR (inter-aural time difference)

ILR (inter-aural level difference)

**latência de comunicação** (delay): Atraso na transferência de dados provocada por comunicações em redes de dados ou computação intensa.

**LAT** (Latency Adaptive Tempo): Função para sistemas de comunicação acústica de rede, cujo tempo musical se adapta, transpondo-se para o valor máximo tolerado pela menor "tolerância à latência" de um músico participante na sessão performativa.

**LAD** (Latency Adaptive Dynamics): Função para sistemas de comunicação acústica de rede, cuja dinâmica musical se adapta, transpondo-se para o valor máximo tolerado pela menor "tolerância à latência" de um músico participante na sessão performativa.

**modulação** (Anel, Frequência, Amplitude): modificação periódica de um parâmetro de um sinal de áudio ou de um processo. Modulação de amplitude e frequência se refere à variação periódica destes parâmetros. Modulação em Anel é um tipo especial de modulação de amplitude onde dois sinais são multiplicados um pelo outro.

MIDI (Musical Interface Digital Instrument): protocolo de transmissão de dados de controle.

**nyquist** (frequência): a frequência limite de sistemas digitais, equivalente à metade da taxa de amostragem. O teorema da amostragem indica que em sinais

digitais, somente frequências situadas entre - e + Nyquist podem ser representadas unicamente. Sinais de frequências fora desta escala serão representadas (erroneamente) por sinais cuja frequencia estará dentro desta escala

**OSC** (Open Sound Control): protocolo de transmissão de dados de controle.

**oscilador**: um componente de sistemas de síntese dedicado a gerar sinais periódicos.

**peer-to-peer**: topologia descentralizada de comunicação em que todos os terminais são em simultâneo clientes e servidores

**ping**: comando de linha que permite medir o tempo que um pacote de dados demora entre o próprio computador (ou instância computacional) e um computador (ou instância computacional) remoto.

**PSO** (Public Sound Objects): sistema de comunicação partilhada em rede para criação coletiva de arte sonora.

**reverberação**: efeito de ambientes (salas, etc.) na projeção, propagação e percepção do som.

ring modulation: modulação em anel.

síntese granular (parâmetros): Tanto a nomenclatura quanto a funcionalidade dos parâmetros de síntese dependem do método adotado. A seguir, fornecemos uma descrição das variáveis comuns à maioria das técnicas analisadas, apontando características específicas quando o conceito não é generalizável. É importante frisar que os parâmetros locais (no nível do grão) podem ser definidos a partir de processos aleatórios, portanto geralmente são utilizados limites inferiores e superiores na hora da implementação.

Grão: unidade elementar das técnicas granulares. As amostras sonoras são definidas pelos seguintes parâmetros: (1) tempo de ataque, (2) taxa de amostragem, (3) amplitude, (4) duração, (5) envelope, e (6) forma de onda.

Tempo de ataque: início do grão em relação ao tempo absoluto.

Taxa de amostragem do grão: Nos paradigmas que usam grãos sintetizados, a taxa de amostragem corresponde à frequência do grão. No processamento granular a partir de uma amostra sonora, a freqüência é expressa por uma razão, onde 1 corresponde à freqüência original do arquivo sonoro. Os valores maiores do que a unidade definem transposições a alturas superiores à da amostra. Os valores menores do que 1 implicam em grãos com alturas mais graves do que altura do som amostrado.

Amplitude do grão: É a amplitude máxima ao longo da duração de um grão. Sendo uma forma de mixagem, a superposição de múltiplos grãos pode exceder os limites do sistema de reprodução sonora. Portanto é aconselhável fornecer um controle global de amplitude do fluxo granular. Também pode ser utilizada uma medida instantânea da amplitude total resultante do processo de granulação (ver *RMS*).

Duração do grão: Intervalo entre o início (tempo de ataque) e o fim do grão.

204

Envelope do grão: O envelope, envoltória ou janela define o perfil de amplitude do grão (cf. envelope e ADSR). Em implementações em tempo diferido, geralmente são utilizados envelopes gaussianos que proporcionam maior definição espectral. A janela trapezoidal aumenta a velocidade e a eficiência dos algoritmos em tempo real (KELLER e ROLFE, 1998). Quando são utilizados envelopes trapezoidais com grãos de duração curta (1 milissegundo, por exemplo), diferentes efeitos espectrais podem ser obtidos alterando a inclinação dos segmentos de ataque e de repouso do envelope (fade-in e fadeout). Na implementação da janela trapezoidal em MacPod (ROLFE e KELLER, 2000), foram utilizados dois tipos de envelope: constante e proporcional. O envelope constante separa a duração do ataque e do repouso da duração do grão. Portanto, os tempos do fade-in e do fade-out são fixos enquanto o segmento de sustentação é variável. Já o envelope proporcional, estabelece o tempo de ataque e repouso a partir da duração granular. Neste caso, a posição da "quebra" do envelope (que controla o Efeito Corner) é constante para todos os grãos (KELLER e ROLFE, 1998).

Forma de onda do grão: É definida pelo conteúdo espectral e temporal do grão. No processamento granular - adotado no modelo PODX - a granulação é realizada acessando amostras sonoras isoladas (TRUAX, 1988). Nesta técnica a forma de onda depende da posição do ponteiro de leitura (pointer) no arquivo de som. Em termos mais gerais, existem quatro maneiras de acessar o conteúdo da amostra sonora: (1) incremental - o arquivo é lido do início ao fim; (2) em loop - o arquivo é lido do início ao fim repetidamente: (3) em ciclo - o arquivo é lido do início ao fim e de trás para frente repetidamente; (4) aleatória - o arquivo é lido a partir de posições aleatórias; e (5) estático — mantendo o ponteiro de leitura numa a posição inicial fixa no arquivo sonoro, o que gera grãos idênticos caso os outros parâmetros locais também permaneçam constantes. Algumas implementações utilizam vários arquivos de som de forma simultânea, podendo combinar diferentes tipos de amostra num único fluxo granular (ROLFE e KELLER, 2000) - ver Barreiro et al. (2009) para uma implementação em que os grãos são amostrados a partir de três arquivos de som pré-gravado. A síntese granular baseada em modelos ecológicos, por sua vez, introduz o conceito de coleção de grãos. A coleção é acessada a partir de parâmetros correlacionados com o evento sonoro sendo modelado (KELLER e TRUAX, 1998). O conceito de coleção de amostras também vêm sendo aplicado em sistemas de mixagem a partir de bases de dados (SCHWARZ, 2006; ZILS e PACHET, 2001), estabelecendo um nexo com os sistemas baseados em dicionários sonoros (STURM et al., 2008).

Estrato granular: Consiste numa seqüência temporal de grãos. "Voz" foi o termo utilizado inicialmente por Truax (1988). No caso da implementação do sistema POD, o uso de memória estava diretamente vinculado ao número de vozes: a mixagem de vinte grãos simultâneos usava a maior parte da memória RAM disponível. Já o uso de CPU dependia da duração dos grãos e do retardo entre grãos – isto é, da taxa de iteração granular - atingindo o limite máximo em durações granulares de aproximadamente um milissegundo. Nos sistemas em tempo diferido, a importância do tempo de processamento é menor. Nesse contexto, o conceito torna-se uma abstração do processo algorítmico que serve para descrever o fenômeno granular. Tendo em vista que o fluxo granular é formado por instâncias múltiplas de grãos acontecendo em paralelo e que os grãos podem ser organizados a partir de processos temporais independentes

(no caso da síntese granular assíncrona) ou correlacionados (na modelagem ecológica), achamos necessário adotar um termo que incorpore os vários tipos de organização do fluxo granular. Retomando as descrições metafóricas de Roads (1991) - como as nuvens granulares — sugerimos, portanto, a adoção do termo estrato granular. O fluxo granular, por sua vez, estaria formado pela sobreposição de estratos múltiplos.

Localização do grão: Posicionamento do grão no espaço virtual representado por três dimensões. Algumas implementações utilizam o paradigma da difusão sonora colocando o fluxo granular em relação ao posicionamento dos altofalantes.

Retardo entre grãos: É o intervalo entre o final de um grão e o início do grão seguinte dentro de um único estrato granular (grain delay).

Sobreposição de grãos: É a medida instantânea da quantidade de grãos ativos.

Taxa absoluta de iteração granular: É definida a partir dos tempos de ataque de grãos consecutivos no mesmo estrato granular (grain rate). O método de síntese granular quase-sincrônica e a síntese por função de onda granular (FOG) utilizam taxas de iteração periódica para obter fluxos granulares com altura definida. Já a modelagem ecológica fornece controle paramétrico da sincronia de fase entre múltiplos estratos granulares. Com esta última técnica é possível criar estratos granulares aperiódicos (sem altura definida) mas com relações de fase fixas, podendo definir um perfil espectral constante ou variável para o fluxo granular.

Taxa relativa de iteração granular: A síntese granular assíncrona utiliza distribuições aleatórias de grãos através do controle do retardo entre grãos intervalo entre o fim e o começo de grãos consecutivos (grain delay). Dessa forma, a somatória do tempo de início e da duração do grão, acrescida ao retardo entre grãos, determina a taxa de iteração granular. A diferença entre utilizar a taxa relativa de iteração e a taxa absoluta de iteração (grain rate) como parâmetros de controle é que esta última permite a manutenção de uma distribuição granular periódica mesmo que a duração dos grãos varie de forma aleatória. Se a duração dos grãos e o retardo entre grãos forem controlados independentemente a partir de distribuições aleatórias, a taxa de iteração será aleatória. Conforme mencionado anteriormente, as distribuições granulares quase-periódicas permitem gerar sons granulares com altura definida. Ao separar o controle da duração do controle do tempo de ataque do grão é disponibilizada mais uma variável para o processo de síntese: o fator de preenchimento (ver mais abaixo).

Sincronia de fase: É a relação temporal entre múltiplos estratos granulares. Um gerador de grãos produz uma série de grãos com parâmetros variáveis ao longo do tempo. O conceito de gerador de grãos implica que apenas um único grão pode estar ativo a cada momento. Assim, quando se deseja ter mais do que um grão simultâneo (grain overlap), vários geradores de grãos devem ser usados. Isso impõe a necessidade de definir a relação de fase entre os diferentes estratos de grãos. A implementação sem sincronia de fase, encontrada nos sistemas de síntese granular assíncrona, produz estratos que são completamente independentes. Quando o tempo entre os grãos nos diferentes estratos precisa ser controlado, uma abordagem fase-síncrona é necessária. Nos

206

modelos ecológicos, a taxa de iteração granular pode ser sincronizada entre os estratos, e a relação de fase pode ser utilizada para estabelecer as características volumétricas ou de profundidade sônica do fluxo granular (TRUAX, 1992) define volume ou profundidade sônica como uma dimensão que engloba as características espectrais e espaciais do som fornecendo pistas em relação à localização e ao tipo de fonte sonora). Portanto, nos fluxos granulares há três configurações possíveis: (1) um único estrato; (2) múltiplos estratos sem sincronia de fase.

Extensão temporal: Expressa como uma razão, a extensão temporal (time-stretching) estabelece o grau de expansão ou contração aplicada ao som original (por exemplo: 0.5 indica metade da duração; 2 indica o dobro da duração). A extensão temporal não se aplica nos casos em que o fluxo granular deriva da utilização de uma coleção de grãos, mas é amplamente utilizado na granulação do som amostrado. A alteração da extensão temporal permite que as características espectrais de um som sejam mantidas enquanto a estrutura temporal do mesmo é alterada (TRUAX, 1992). Portanto, através da expansão temporal é possível salientar certas características internas do som, o que resulta numa espécie de microscópio sonoro (ou lupa) dentro do contexto do trabalho em paisagem sonora (soundscape) e configura um exemplo de processamento mais do que de síntese - o conceito de extensão abrange tanto a expansão quanto a contração temporal da amostra sonora. O termo em inglês stretch configura um caso de marcação lingüística (COELHO DE SOUZA, 2010).

Densidade granular: É uma propriedade emergente da interação entre a taxa de iteração granular e o número de estratos. A densidade pode ser controlada de diversas formas: utilizando parâmetros globais (pelo número de estratos em síntese granular assíncrona ou pela quantidade de energia inserida no sistema de ativação na modelagem ecológica) ou a partir de parâmetros locais (como o retardo entre grãos na síntese assíncrona ou a taxa de interação na síntese por função de onda granular).

Fator de preenchimento: É a proporção entre a somatória das durações granulares e a duração total do estrato granular. O fator de preenchimento está diretamente relacionado ao retardo entre grãos, mas é independente dos tempos de ataque e da quantidade de grãos (ver KELLER e BERGER, 2001). Portanto, fornece um parâmetro complementar à densidade granular.

Síntese por tabela de onda: neste metodo, uma onda sonora é colocada em um bloco de memória e acessada continuamente (geralmente por um oscilador) para a geração de sinais.

**sample theorem**: teorema da amostragem (ver *nyquist*, *frequencia*).

sampling: (ver amostragem).

**som**: ondas (normalmente longitudinais) de compressão e expansão do meio propagante (normalmente o ar).

**som** (velocidade de propagação): 340,29 m/s no ar, em condições normais de pressão e umidade.

sones: curvas de iso-loudness.

