

TRILHA SONORA DA INTERFACE: DIÁLOGO, AMBIÊNCIA E EFEITOS SONOROS

INTERFACE SOUNDTRACK: DIALOGUE, AMBIENCE, AND SOUND EFFECTS

Luiz Roberto Carvalho¹, M. Sc.
Alice Theresinha Cybis Pereira², PhD.

(1) *Universidade Federal de Santa Catarina (UFSC)*
e-mail: semprecarvalho@gmail.com

(2) *Universidade Federal de Santa Catarina (UFSC)*
e-mail: acybis@gmail.com

Design de Interface, Trilha Sonora, Áudio Dinâmico

O conjunto dos sons percebidos pelo ouvido humano numa apresentação multimídia pode ser definido como trilha sonora. Tradicionalmente utilizada no design de som de filmes e de games, a trilha sonora possui três categorias: diálogo, ambiência, e efeitos sonoros. Nos ambientes interativos, mais especificamente no ambiente web, a aplicação dos conceitos pertinentes à trilha sonora encontra dificuldades, tendo em vista a vasta subutilização e desvalorização dos elementos sonoros nessas plataformas. Para que se possa iniciar o desenvolvimento de um entendimento formal dos parâmetros que constituem o design de som nas IHC, o presente estudo objetiva expor as principais aplicações do som na interface, no que diz respeito às suas características de conteúdo, forma e função, que são manifestados através da utilização de diálogo, ambiência e efeitos sonoros. Dessa forma, busca-se sistematizar e simplificar o processo de criação, produção e implementação de sons nas interfaces.

Interface Design, Soundtrack, Dynamic Audio

The set of sounds perceived by the human ear in a multimedia presentation can be defined as a soundtrack. Traditionally used in sound design for movies and games, the soundtrack has three categories: dialogue, ambience, and sound effects. In interactive environments, more specifically in the web environment, the application of concepts pertinent to the soundtrack encounters difficulties, due to the vast underutilization and devaluation of the sound elements in these platforms. In order to begin the development of a formal understanding of the parameters that constitute sound design in HCI, this study aims to expose the main applications of sound in interfaces, due to its characteristics of content, form and function, which are manifested through the use of dialogue, ambience and sound effects. This approach seeks to systematize and simplify the process of creation, production and implementation of sounds in interfaces.

1 Trilha Sonora da Interface: Diálogo, Ambiência e Efeitos Sonoros

[SONNENSCHNEIDER, 2001], [HUG, 2011] e [JØRGENSEN, 2011] definem o conjunto dos sons percebidos pelo ouvido humano numa apresentação multimídia como trilha sonora (do inglês *soundtrack*; tradução livre), e apontam três categorias, nomeadamente diálogo, ambiência, e efeitos sonoros, como as tradicionalmente utilizadas em design de som de filmes e *games*. Nesse contexto, [SERAFIN, 2004], [FOLLETT, 2007], [MOSES, 2010] e [WILHELMSSON e WALLÉN, 2011] apontam que os eventos acústicos ocorridos numa interface podem ser divididos do mesmo modo, em:

- a) Diálogo, representando os sons que expressam a linguagem falada;
- b) Ambiência, formada pelo som ambiente e música;
- c) Efeitos sonoros, correspondentes aos *auditory icons* e *earcons*.

De maneira geral, a trilha sonora fornece contexto para o estabelecimento de uma linguagem narrativa na interface. [FOLLETT, 2007] assinala que o som ambiente constrói atmosferas, temas musicais ajudam a estabelecer uma relação de ordem emocional com a interface, e os efeitos sonoros orientam os usuários para determinados pontos da interface, notificando ocorrências que não podem ser visualizadas e atuando como elementos de sinalização e orientação. [COLLINS, 2008, p. 78] corrobora com tal afirmação ao elucidar que:

“Do ponto de vista cognitivo, o canal auditivo é uma excelente maneira de transmitir informações e dar *feedback* ao usuário. Enquanto a música de fundo pode ser utilizada para evocar emoções e definir um ambiente, um efeito de som específico pode ser utilizado para transmitir informações específicas de alerta”.

[LILJEDAHL, 2011] salienta que é importante mencionar que as categorias de diálogo, ambiência e efeitos sonoros são pontuadas apenas para fins de um melhor entendimento das mesmas, e suas

possíveis fronteiras estão em constante intercâmbio. A música já é utilizada como parte integrante dos efeitos sonoros, bem como efeitos sonoros são utilizados como trechos de músicas, entre outras variações. De maneira geral, a utilização da trilha sonora em ambientes interativos oferece uma consistente estrutura para o alcance de acentuados níveis de imersão, no entanto, ainda observa-se uma relativa ausência de ferramentas de análise com capacidade suficiente para extrair informações funcionais e estéticas relativas ao impacto do som nos sistemas interativos.

2 Diálogo

De acordo com [LILJEDAHL, 2011], quando da necessidade de transmitir mensagens claras e inequívocas, a voz humana é uma escolha natural. Dentre os três tipos de trilha sonora, a voz é a linguagem sonora mais facilmente assimilada pelos indivíduos, sendo geralmente utilizada para transmitir mensagens claras e com reduzido risco de interpretação equivocada. Como parte integrante e indissociável da formação da própria sociedade, a voz é capaz de reduzir a distância entre a interface e o usuário, podendo ser utilizada para promover concentração e imersão através do canal auditivo.

Nos diálogos, a fala humana pretende ser informativa. Ao se expressar por meio da fala, há uma explícita atenção nas palavras que são proferidas, mas adere-se também, implicitamente, à forma como essas palavras são ditas. Assim, o significado da fala não é exercido unicamente pelo conteúdo verbal, mas também pelas qualidades vocais que são impostas durante o discurso. Nesse sentido, todas as variações de qualidades subjetivas, expressivas e dramáticas da voz humana podem ser utilizadas na interface. Uma sentença pronunciada com tom agressivo e “áspero” será interpretada de modo distinto da mesma frase pronunciada com uma voz “amiga” e “solidária”. [CANCELLARO, 2006] pontua que o mais importante requerimento para o uso da voz humana é que esta deve ser a mais límpida, clara e audível, e nenhum outro tipo de som deve estar ocupando a mesma faixa de frequência com ênfase quando a voz humana é disparada na hipermídia.

[VAN LEEUWEN, 1999] dividiu e classificou a voz humana em distintas propriedades, que carregam significados comunicativos culturalmente formados, a saber:

- a) tensão: quão firme ou tensa a voz pode ser;
- b) rugosidade: quão rouca e gutural (com tons mais ásperos, sendo mais associados com homens);
- c) *arejamento*: quão arejada ou íntima (nas culturas ocidentais, geralmente vozes mais arejadas são tidas como menos autoritárias);
- d) sonoridade: quão expansiva ou suave;
- e) timbre: quão alta ou baixa (em termos de espectro de frequências, o quão aguda ou grave esta voz pode ser);
- f) vibração: qual o nível de tensão (em termos de vibração de cordas vocais, associada ao termo técnico *vibrato*¹).

Na realidade cotidiana, a linguagem falada é usada principalmente de maneira informal. Os significados das palavras e frases não são fixos por definição, mas dependem de um contexto formado pela sua definição geral, pela interação conversacional, e de acordo com a sequência de sentenças proferidas, somado ao uso de metáforas e analogias. Neste contexto reside a grande importância do adequado uso da linguagem verbal na interface, de tal modo que essas características intrínsecas da comunicação verbal possam também estar presentes nos ambientes interativos. Um cuidado a ser tomado em relação ao diálogo é o risco deste tornar-se repetitivo, caso o usuário já o tenha ouvido anteriormente, e não necessite ouvir novamente. Nesse caso, [LILJEDAHN, 2011] recomenda permitir aos usuários a opção de ignorá-los quando estes não se fazem mais necessários.

¹ Vibrato é um efeito musical que consiste em uma mudança regular e pulsante de afinação. É utilizado para adicionar expressão à música vocal e instrumental. Ao cantar, ocorre espontaneamente através de um tremor nervoso no diafragma ou laringe [SUNDBERG, 1994].

2.1 Comandos de Entrada e Saída por Voz

Quando algoritmos são usados para gerar sons, como ocorre com o mecanismo de geração de voz (*TTS*, do inglês *text-to-speech*), a entonação pode soar robótica (ou não humana), porém, para algumas aplicações, o som robotizado de computador pode ser preferido, como apontado por [SHNEIDERMAN e PLAISANT, 2009, p. 382; *tradução livre*], indicando que “os sons de robô utilizados no metrô do aeroporto de *Atlanta* chamaram mais atenção do que a gravação de um ser humano, quando da tarefa de orientar as direções”. Quando as mensagens são simples e curtas, os segmentos de fala digitalizada são geralmente preferíveis. O recurso torna-se particularmente vantajoso para os usuários quando seus canais visuais estão sobrecarregados, e estes devem estar livres para se movimentar, ou quando o ambiente é inapropriado para a utilização de uma interface gráfica. No entanto, o imediatismo no impacto emocional de uma conversa telefônica é um componente convincente da importância da comunicação entre seres humanos, e da necessidade de uma *humanização* dos ambientes interativos.

[KAUSHANSKY, 2012] descreve um experimento que opôs uma boneca *Barbie* e um *Furby* (popular brinquedo peludo, que conversa com os usuários) para verificar por quanto tempo crianças realizariam a tarefa de deixá-los pendurados de cabeça para baixo. Todas as cinco crianças no experimento foram capazes de segurar a *Barbie* de cabeça para baixo sem ressalvas, no entanto, o brinquedo de conversa eletrônica não foi tratado da mesma forma, pois assim que este era colocado de cabeça para baixo um feedback sonoro era disparado com os dizeres “me assustou”, conferindo características emotivas ao brinquedo. Após o experimento, as crianças relataram que não queriam que o brinquedo “ficasse com medo”. Nesse sentido, foi possível concluir que a componente verbal humaniza os dispositivos, pois envolve o usuário e cria uma ligação emocional com o mesmo. O programa *Siri* da *Apple* (2015), dispositivo de interação por meio da voz, já demonstra uma crescente popularidade dos dispositivos de troca de informação através da voz.



O programa já é capaz de catalogar preferências, compromissos, atividades diárias, e sugerir e antecipar solicitações feitas anteriormente pelo usuário. Após detectar o nome do usuário, *Siri* passa a utilizá-lo em suas respostas, adicionando assim uma conexão mais íntima e pessoal na interação com os produtos da *Apple*.

Menus de áudio também são importantes em situações de interfaces de acesso público, que precisam acomodar usuários cegos ou deficientes visuais, como no caso de quiosques de informação ou máquinas de votação. *Smartphones* com sistemas *Android* e *iPhones* já possuem opções de acessibilidade que permitem experiências com entrada gestual e de áudio e saída de áudio que vem sendo amplamente utilizadas por deficientes visuais.

3 Ambiência: Som Ambiente e Música

O termo ambiência é utilizado para descrever a sensação auditiva e subjetiva que faz o ouvinte perceber o tamanho e demais características de um ambiente ou recinto em que foi capturado determinado som. Capazes de criar uma sensação de presença física, os sons ambientes criam um senso de atmosfera e presença no ouvinte, e são capazes de estabelecer tons de humor e tensão. Em filmes, por exemplo, se um casal está caminhando na praia e um som ambiente de ondas é colocado estrategicamente, a atmosfera oceânica é intensificada. Na interface, a ambiência pode ser formada pela utilização de sons ambientes, pela inserção de uma peça musical, ou pela junção destes dois elementos. Na hiperfídia, a ambiência define o clima e o ritmo da narrativa da interface e está ligada a interpretações de cunho emocional, podendo despertar reações distintas no usuário. [PARKER e HEEREMA, 2008] definem de *background music* (música de fundo, em tradução livre) como a trilha de áudio contínua, sem pausas, que está presente na interface para que o usuário tenha uma sensação de presença e, assim, sintam-se imerso.

[LILJEDAHN, 2011] afirma que mesmo que por vezes praticamente inaudível, e por mais simples e discreto que possa soar um som ambiente, ao

adicioná-lo na interface, torna-se possível criar uma experiência imediata de presença e realidade. No entanto, [FOLLETT, 2007] pontua que enquanto o som ambiente funciona bem no design de games para estabelecer tons de humor e localização espacial, ele não possui um papel claro em outras experiências de usuário, como no caso de aplicações de software e nas interfaces *web*. De maneira geral, a ambiência é um elemento-chave para a definição de “climas” (ou atmosferas), podendo incluir em sua constituição efeitos sonoros, composições musicais, e sons ambientes propriamente ditos. A criação de climas pode ser utilizada para antecipar ao usuário um contexto específico, ou informá-lo acerca de mudanças que estão ocorrendo na interface. Como já ocorre nos games, o designer de som Steve Kutay elucidou:

“O impacto psicológico de sons ambientes pode acrescentar muito à imagem na tela, embora não visualmente presente no cenário. Por exemplo, um grito distante sustentado por uma criança sugere vulnerabilidade ou insegurança. Uma cerca quebrada chacoalhando com o vento de uma cidade abandonada sugere ao leitor um evento anterior, talvez traumático. Estes são exemplos sutis usados para despertar a consciência do usuário. Sons mais óbvios devem ser usados para alertar o jogador da proximidade direta com o perigo: atmosferas escuras ou vocalizações abafadas de um possível inimigo irão preparar o jogador para um combate feroz à frente. Antecipação, medo e ansiedade são facilmente evocados pela colocação cuidadosa de sons ambientes” [KUTAY, 2006; tradução livre].

3.1 Música

Dentre todos os itens que compõem a trilha sonora, a música pode ser considerada um dos elementos chave para estabelecer induções de humor. [Liljedahl, 2011] pontua que a música é parte integrante da cultura humana desde os primórdios do *Homo sapiens*, e é muitas vezes referida como a linguagem da emoção, pois é capaz de evocar uma diversidade de sentimentos, como alegria, tranquilidade, agitação, tristeza. A música é categorizada não somente por sua estrutura, mas pela abstração de eventos e experiências vividas pelo indivíduo que a escuta, de acordo com fenômenos culturais, sociais ou psicológicos, que vão desde situações particulares de consumo musical, de condutas corporais ou sociais

produzidas em torno da música, de processos subjetivos, de relações interpessoais e coletivas, entre tantos outros elementos que compõem o repertório de experiências vividas em torno da música. [ILARI, 2006] aponta que a música é um fenômeno social que vem mantendo funções tradicionais e sentidos próprios em diferentes sociedades no decorrer da história. No mundo ocidental, a música vem exercendo funções específicas em atividades humanas como dançar, contar estórias, comemorar datas especiais, rezar, entreter, vender produtos, e estas e muitas outras funções da música na vida cotidiana estão claramente relacionadas às relações interpessoais. A autora ilustra o exemplo dos Iorubás da África, em que o uso da música implica na ideia de parentesco, religião, política e economia, e dos capoeiras, que associam a música ao movimento corporal, ao ritual e à libertação.

Nesse sentido, tanto as funções quanto os significados do fazer musical dependem de aspectos específicos de cada sociedade e cultura. Enquanto uma peça de música rápida e em volume alto pode ser usada para elevar os níveis de excitação durante uma festa, uma canção suave e serena pode servir para acalmar uma criança. As formas de utilização e apreciação da música variam de acordo com uma infinidade de combinações de crenças pessoais e objetivos, sendo que o senso de “música apropriada” depende de fatores culturais e situacionais, uma vez que as percepções e usos da música são comportamentos aprendidos e previamente determinados por membros de um grupo social específico.

A música faz referência a toda uma variedade de objetos acústicos e não acústicos, podendo ser associada a outras formas de expressão, pertencentes aos domínios da visualidade e do verbal, como aponta [CANCELLARO, 2006], ao afirmar que a música é capaz de influenciar uma experiência visual. Para Martinez (2000), tais formas, como a dança, o cinema, a multimídia, e a própria hiperfídia, constituem linguagens que canalizam suas significações através dos dois principais sentidos humanos, a audição e a visão – os principais sentidos estéticos. A eficiência estética destas formas de expressão, que fazem da

música sua base ou que compartilham com ela suas propriedades e estruturas, está na possibilidade de apresentação e processamento de signos híbridos, tendo em vista que através do endereçamento dos dois principais sentidos humanos é possível estabelecer acentuados níveis cognitivos.

[COLLINS, 2011] aponta que a música é um dos elementos-chave que impulsionam a emoção, e que a adição de elementos musicais nas interfaces pode proporcionar um maior envolvimento do usuário, sendo possível obter distintos resultados, desde gerar um momento dramático, para chamar a atenção, até tranquilizar os usuários com composições calmas e relaxantes. A presença de música nas interfaces, independente de seu conteúdo, indica de que o que está sendo experimentando é distinto das demais seções que não possuem música. [STEVENS e RAYBOULD, 2011] afirmam que nos games a música muitas vezes marca o significado de eventos especiais, como o início, a execução e a conclusão de tarefas ou seções da interface. Ao utilizar a música de maneira estratégica e pontual, é possível aumentar o impacto e a eficácia destes pontos específicos. No entanto, se a música estiver presente durante todo o percurso de navegação do usuário, seu impacto será reduzido.

A música é um fenômeno linear: possui início e fim delimitados, e a jornada entre os dois contém a mesma quantidade de tempo. Isto se aplica verdadeiramente tratando-se de composições gravadas. Em sua maioria, as composições musicais possuem um formato que cria sucessões de tensão e alívio, que por sua vez, criam expectativas sobre como a composição irá continuar. Nesse sentido, [LILJEDAHN, 2011] aponta que para funcionar e ser percebida como música, a composição deve possuir temas indicativos de forma e continuidade definidas. No entanto, no contexto das mídias interativas, que são essencialmente não lineares, o conceito de música linear não se aplica adequadamente, tendo em vista que a música não possui uma estrutura que permite uma imediata adaptação de acordo com as decisões advindas do usuário ou do sistema. A maioria das questões decorrentes da utilização adequada da música em interfaces surge do conflito entre a



natureza interativa do sistema, onde eventos podem acontecer a qualquer momento, e a componente temporal da música, que se baseia na linearidade e em comprimentos predeterminados de trechos musicais.

3.2 Variabilidade da Música Interativa

Para que a música possa ser utilizada com eficácia em ambientes interativos, duas questões fundamentais devem ser consideradas: as distintas interfaces gráficas e as variações dos trechos musicais. Se a música irá responder aos estímulos do usuário, as ações de navegação deverão ser correspondentes às transições entre os trechos musicais, logo, a troca de estados musicais e de um estado musical para um não musical não deve ser feita de maneira linear e abrupta. A indústria de games têm desenvolvido sistemas para lidar com a distinção entre a natureza linear da música e a dinamicidade dos ambientes interativos, tornando possível com que as composições musicais modifiquem suas estruturas de acordo com as mudanças realizadas pelos usuários e pelo ambiente. [STEVENS e RAYBOULD, 2011] pontuam a importância dos testes para a utilização da música nos games, tendo em vista que o grau de reatividade ou interatividade da música só pode ser estabelecido através de testes e feedback completo.

De acordo com [KOJI KONDO, 2007], compositor musical da série *Super Mario*, a música deve evidenciar o caráter participativo da interface, podendo indicar mudanças de estado com alterações no ritmo das composições, adição de instrumentos e vozes, e até mesmo alterando a música de acordo com distintos estágios de interface, com a possibilidade de adição ou variabilidade na reprodução de sentenças sonoras. Nesse sentido, a música dinâmica possui distintos níveis nos quais deve reagir ou interagir, sempre de acordo com a narrativa, e em resposta direta ao usuário. O autor descreve quatro componentes da música dinâmica, típica de ambientes interativos:

1. A habilidade de criar músicas que mudam de acordo com a interação do usuário;
2. A habilidade de criar uma produção

multicolorida; criar distintos temas para a mesma composição;

3. A habilidade de adicionar elementos-surpresa nas composições musicais, de modo a intensificar a interatividade;

4. A habilidade de adicionar elementos musicais como características específicas de estado de interface.

Existem várias abordagens para que se possa executar uma variação musical, de modo que uma composição se estenda e ao mesmo tempo possa ser mais sensível aos *inputs* do usuário e à narrativa. Um compositor pode compor centenas de variações para as pistas de áudio, ou pode introduzir algum tipo de variabilidade na mesma composição, de acordo com as seguintes abordagens apontadas por [COLLINS, 2008]:

- a) Tempo Variável: ajustes de tempo são bastante comuns em games, como já ocorrido a partir do final dos anos 1970, com o título *Space Invaders* ou *Asteroids*, sendo que ambos títulos possuíam sons que aceleravam seu andamento conforme o jogo progredia. Geralmente, este conceito tem sido usado em relação a questões temporais na interface: o usuário possui determinado tempo para cumprir com uma tarefa específica do aplicativo, e assim que este tempo está esgotando-se, o ritmo da música acelera-se, como ocorre, com o game *Super Mario Bros*, por exemplo. Porém, deve-se estar atento para que as variações de tempo não tornem a aplicação *cartunesca*, tendo em vista a utilização exaustiva desses recursos por boa parte das animações infantis, em particular os clássicos humorísticos de *Walt Disney*, como o *Mickey Mouse*, por exemplo;

- b) Volume Dinâmico Variável: simples variações de volume são bastante comuns em games, principalmente nas interfaces gráficas de menus, nas quais o volume é reduzido enquanto o jogador faz seleções, antes de retornar ao volume total durante a ação de jogabilidade. Reduzir o volume completamente também é uma opção quando o usuário passou muito tempo em determinado ambiente de interface. Assim que um novo padrão

de interação surge, algo novo ocorre, o volume da música pode retornar novamente. Outros parâmetros, como tensão de narrativa, podem também ter relação com o volume da interface. Em momentos de calma e concentração, os sons podem ser reproduzidos em volumes reduzidos, mas à medida que a tensão e ação requeridas aumentam, os sons podem aumentar sua intensidade exponencialmente;

c) Mixagem Variável: mudanças na instrumentação ou orquestração também podem oferecer soluções interessantes para a música dinâmica. Ao ajustar vários instrumentos num cenário sonoro - adicionando percussão, por exemplo, para sugerir tensão -, um estado de humor pode ser introduzido na narrativa. É possível ocultar elementos no panorama de mixagem, e apenas trazê-los novamente em momentos importantes, em camadas. Nesse sentido, a música é composta em camadas, com distintos instrumentos e vozes, que podem ser removidos ou adicionamos a qualquer momento.

Devem-se usar tais recursos de modo que a composição musical esteja intrinsecamente relacionada à narrativa e aos demais elementos da interface. Se a composição musical não possui uma curva de tensão e não há um clímax, esta pode facilmente tornar-se um mero papel de parede sonoro, atuando apenas como música ambiente. Como tal, a peça sonora perde suas funções em relação à narrativa. [COLLINS, 2011] assevera que a música na interface deve ser tão comunicativa quanto possível, e sua atuação não deve levar a uma falta de significação. Seus elementos variáveis devem ser determinados de acordo com objetivos estabelecidos, apresentando assim uma definição, uma caracterização, um sentido de desenvolvimento suficiente para sua aplicação. Nos games, transições importantes de narrativa ocorrem geralmente em momentos de tensão, momentos em que a música é fundamental para enfatizar a ação do usuário.

4 Efeitos Sonoros

Por definição, efeitos sonoros são disparos de som breves e têm como função destacar algum aspecto

da interface. [SHNEIDERMAN e PLAISANT, 2009] assinalam que desde a origem das interfaces *desktop*, uma série de sons foi utilizada na indicação de tarefas, avisos, ou como forma de reconhecimento da conclusão de uma ação, como os clássicos disparos sonoros indicativos de arquivos sendo enviados à lixeira, ou de recebimento de e-mails. Quando se trata de design de som para UX, [FOLLETT, 2007] pontua que a utilização do som se dá na sua grande maioria através dos efeitos sonoros. A presença de efeitos sonoros de áudio interativo-direto em botões, barras de rolagem e ações de usuário em sites é mais frequente do que os demais elementos da trilha sonora da interface, como diálogo, ambiência e música.

Os efeitos sonoros podem ser representativos dos sons dos objetos que fazem parte do universo físico, na medida em que os objetos representados metaforicamente na interface assemelham-se aos objetos reais. Nesse sentido, [LILJEDAHN, 2011] pontua que um retângulo movendo-se horizontalmente em silêncio no monitor de vídeo é provavelmente visto como apenas um retângulo na tela. No entanto, se o som de uma pesada pedra sendo arrastada for adicionado nesta animação, o retângulo passa a ser considerado um elemento com propriedades físicas, como peso, densidade, similares às características da pedra pesada. Nesse sentido, [KAUSHANSKY, 2012b] pontua que é possível incorporar um significado de ação específico através do uso de um único efeito sonoro e desse modo oferecer um feedback rápido e eficiente ao usuário. Como forma de obter a atenção do usuário de ambientes interativos, [WICKENS *et al*, 2012] sugerem a utilização de sons suaves como ferramenta de feedback positivo e de frequência regular, e de sons fortes para situações de emergência, que possuem frequência reduzida, e onde há a necessidade de pronta tomada de decisão.

[PARKER e HEEREMA, 2008] pontuam que os efeitos sonoros possuem múltiplas funções, e enfatizam que o disparo sonoro de confirmação de ação é o mais comumente utilizado. Segundo [BREWSTER, 1994], um dos pioneiros nos estudos de som no domínio da IHC, o mais

tradicional uso de efeitos sonoros na interface se dá através da utilização de alarmes e advertências ou para monitoramento de tarefas. Alarmes são sinais projetados para interromper uma tarefa em curso e indicar que algo requer atenção e ação imediata, já o monitoramento objetivo fornecer informações sobre alguma tarefa em curso. [SHNEIDERMAN e PLAISANT, 2009] pontuam que efeitos sonoros podem fornecer *feedback* informativo sobre o progresso de determinada tarefa, bem como alertar os usuários em situações de emergência. No entanto, um mecanismo para suprimir alertas deve ser fornecido, e se vários tipos de alarmes são usados, é necessário testá-los para garantir que os usuários irão distinguir as ações cabíveis a cada nível.

Porém, o uso interativo de som deve ser visto como distinto das aplicações que servem primariamente para sinalizar ou alertar um usuário a um evento, processo ou estado. [ROCCHESSO *et al*, 2008] apresentam algumas formas em que o som pode assumir um papel mais proeminente: a) o som pode criar ou revelar novas funcionalidades em um produto, quando estas funcionalidades podem não ser aparentes através de outras modalidades, tais como a forma física ou outros indicadores visuais; b) o som pode também ser usado para fornecer *feedback* para ajudar o desempenho dos usuários através de uma interface, como por exemplo, uma ferramenta de controle cardíaco na execução de atividades físicas.

Os efeitos sonoros são mais curtos do que avisos verbais, e podem ser menos intrusivos. No entanto, há uma carga limitada de informação que pode ser incorporada num estímulo sonoro curto, sendo o uso da linguagem verbal mais apropriado para a transmissão de mensagens complexas. Em situações menos críticas das utilizadas em alarmes, é possível utilizar um lembrete de voz para indicar que um usuário tenha perdido um campo quando o preenchimento de um formulário. [FOLLETT, 2007; *tradução livre*] exemplifica ao pontuar que em vez de forçar um usuário a visualizar uma página de formulário inteira para encontrar o texto em vermelho indicativo da falta de preenchimento do mesmo, uma voz poderia simplesmente dizer “por favor, digite seu nome”.

Existem distintas formas para apresentar informações através da utilização de efeitos sonoros, no entanto, há duas categorias que vem sendo estudadas por mais de duas décadas: os ícones auditivos, ou *auditory icons* [GAVER, 1986; 1989; 1992; 1993; 1993b] e os ícones sonoros, ou *eacons* [BLATTNER *et al*, 1989]. [SHNEIDERMAN e PLAISANT, 2009] apontam uma distinção útil de sons dinâmicos e os classificam em sons familiares, referenciando os ícones auditivos (do inglês *aditory icons*), e sons abstratos, em referência aos ícones sonoros (do inglês *earcons*).

Ícones auditivos são sons familiares, como o som de uma porta se abrindo, uma bola quicando, e ajudam a reforçar as metáforas² visuais da interface. Ícones sonoros são sons abstratos, dizem respeito aos sons desenvolvidos especificamente no domínio da interface, sons que não possuem nenhum tipo de associação externa ao seu significado. Nesse sentido, os efeitos sonoros podem ser criados através da captura de sons advindos do ambiente acústico, possuindo características do ambiente, ou pode ser artificial, sendo o resultado de um processamento eletrônico de sinal digital.

Outras categorias de uso de efeitos sonoros utilizados na interface descritos por

² Segundo [SANTOS, 2009, p.109], “metáforas, no contexto computacional, são associações entre o mundo do usuário e conceitos do universo computacional”. São utilizadas extensivamente devido ao poder de familiarização que proporcionam, como por exemplo, a clássica metáfora da mesa de escritório, com pastas, fichários, lata de lixo, presentes na interface gráfica dos sistemas operacionais. Categorizações classificam as metáforas como concretas (um objeto, como um livro) ou conceituais (um sentimento, como o amor), sendo que as concretas são as mais apropriadas para usuários iniciantes e casuais, bem como são as que atendem a um público mais amplo, pois se utilizam de conceitos mais familiares. Isto se dá partindo-se da premissa de que uma interface com objetos já previamente conhecidos pelos usuários possui uma tendência positiva para a assimilação de seu uso, ou seja, se o usuário já está familiarizado com os conceitos associados às metáforas, este será capaz de transpor essa experiência prévia para o domínio da interface.

[SHNEIDERMAN e PLAISANT, 2009] são os sons *cartunescos* (ou *cartoonified*), que exageram os aspectos de sons familiares. Sons *cartunescos* devem ser considerados como também pertencentes ao grupo dos sons abstratos, por não possuírem – em sua maior parte – relação com os sons percebidos pelo ambiente circundante dos indivíduos, dadas suas características de exagero e aumento de proporções. A figura abaixo esquematiza uma divisão com base nos conceitos propostos pelo autor:

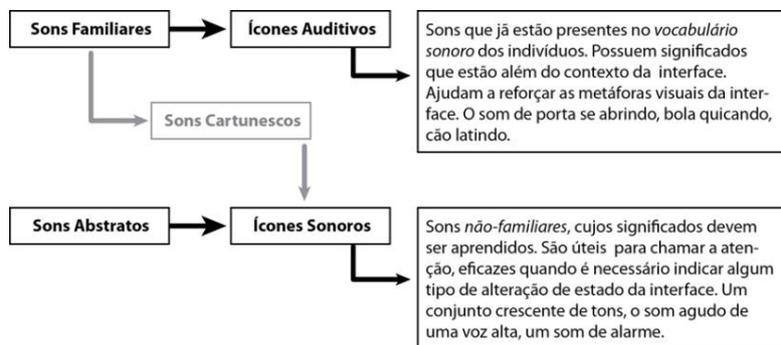


Figura 1: Divisão de sons na interface: familiares, abstratos e cartunescos. Fonte: CARVALHO & PEREIRA, 2012, p. 255.

4.1 Ícones Auditivos; Auditory Icons

[GAVER, 1986; 1989; 1992; 1993; 1993b] foi quem desenvolveu o conceito de ícones auditivos, que diz respeito à utilização de sons que já são previamente conhecidos pelos usuários, que já fazem parte de seu repertório auditivo. Trata-se de sons naturais, que os indivíduos ouvem com frequência, sons de eventos que são capturados em ambientes físicos, como por exemplo, o som de bater a palma das mãos, que é utilizado para representar uma ação ou objeto na interface. Ao ouvir o som de uma porta batendo, o ouvinte é capaz de prever o tamanho, o material da porta, a quantidade aproximada de força que foi utilizada, bem como o tamanho da sala em que a porta está instalada. Essas características advindas das trocas físicas entre objetos e suas resultantes acústicas podem ser utilizadas na interface. Ao selecionar um objeto numa interface, um som pode ser disparado de acordo com o tipo de objeto selecionado, com o tamanho do arquivo que

representa o objeto, entre outras características relevantes que podem ser expostas, partindo-se da relação entre as propriedades físicas dos objetos e suas componentes acústicas.

Os sons têm uma ligação intuitiva com aquilo que eles representam, por exemplo, ao selecionar um ícone, um som de toque entre duas superfícies pode ser disparado, tendo em vista que o usuário pressiona no ícone com o cursor sobre o mesmo. No entanto, problemas podem ocorrer com sistemas representacionais de ícones auditivos, tendo em vista a natureza abstrata de algumas ações de interface, dada à existência de objetos virtuais que não possuem representação acústica compatível com algo externo à própria interface. A fragilidade dos ícones auditivos reside no fato de que algumas situações na interface não possuem equivalentes aos sons experienciados no universo físico, como por exemplo, não há equivalente sonoro para a pesquisa em um banco de dados, logo, sons reais podem não ser representativos da ação na interface.

[GAVER, 1993] pontua que os ícones auditivos são adequados para o transporte de informação semântica sobre eventos. Em seus resultados, o autor aponta que os indivíduos não esquecem os significados dos sons, sendo que uma das maiores vantagens dos ícones auditivos é a capacidade de comunicar significados que os ouvintes podem facilmente aprender e lembrar, diferentemente de outros sistemas que utilizam sons abstratos, onde os significados são mais difíceis de aprender, como ocorre com os ícones sonoros, ou *earcons*.

4.2 Ícones Sonoros; Earcons

Desenvolvidos por [SUMIKAWA, 1985], [SUMIKAWA et al, 1986], e [BLATTNER et al, 1989], ícones sonoros, ou *earcons*, são compostos através da formação de timbres advindos de sons abstratos pertencentes ao domínio digital, denominados sons sintetizados, que são moldados a partir de combinações estruturadas para criar mensagens sonoras e representar porções da interface. [BLATTNER et al, 1989] definem como ícones sonoros todas as mensagens de áudio não-verbais que são utilizadas na interface para

fornecer informações ao usuário sobre algum objeto do universo virtual. Ao contrário dos ícones auditivos de [GAVER, 1986; 1989; 1992; 1993; 1993b], não há uma relação intuitiva entre o som e o que ele representa, e esta associação deve ser aprendida pelo ouvinte.

Ícones sonoros são construídos a partir de blocos denominados *motivos*, que são pequenas sub-unidades de sequências rítmicas que podem ser combinadas de maneiras distintas. [SUMIKAWA et al, 1986] descreve o motivo como uma breve sucessão de valores musicais dispostos de tal forma a produzir um padrão de tons suficientemente distinto para permitir que este funcione como uma entidade reconhecível individualmente. A eloquência dos motivos reside na sua capacidade de serem combinados para criar estruturas reconhecíveis maiores, sendo que a repetição de motivos ou a interconexão de motivos distintos produz padrões autossuficientes maiores. [SUMIKAWA, 1985] pontua o ritmo e a afinação como parâmetros fixos de ícones sonoros, e o timbre e a dinâmica como parâmetros variáveis. Os parâmetros fixos definem um motivo, e os variáveis são capazes de alterar suas características. O autor aponta três possíveis formas de manipulação de motivos para a criação de ícones sonoros: a repetição, a variação e o contraste.

5 Considerações Finais

Ao pontuar os elementos integrantes da trilha sonora dos ambientes interativos, oferece-se um adequado suporte conceitual capaz de embasar uma aproximação técnica, e prática, acerca da utilização do áudio na interface.

Nesse sentido, é pertinente iniciar o entendimento formal dos parâmetros que constituem o design de som nas IHC, de modo que tais estudos possam sistematizar e simplificar o processo de criação, produção e implementação de sons em ambientes interativos. É necessário não somente avançar com a discussão acerca dos papéis exercidos pelo som na interface, mas também assinalar a necessidade de uma nova pedagogia de design de interface que evidencie o importante papel exercido pelo som no

campo das atividades curriculares acadêmicas.

Há uma escassez de terminologias e metodologias que possam embasar uma discussão mais aprofundada em relação à interatividade e o som. Nesse sentido, a ausência de parâmetros de utilização do som nas interfaces, tanto na tradicional interface *web* quanto nas aplicações direcionadas aos dispositivos portáteis, como *tablets* e *smartphones*, pode resultar na utilização de sons intrusivos e indesejados, e comprometer a experiência do usuário. Trabalhos com uma consciente utilização do som na área de design de interface ainda encontram-se num período tão nascente que há uma grande necessidade em explorá-los. No entanto, a apropriação do canal de áudio em experiências interativas não deve ser confundida com a mera colocação de sons sobre os elementos gráficos. Desde sua origem, sons têm sido utilizados de maneira *ad hoc*³ para cada interface, resultando em experiências muitas vezes frustradas, seja porque os próprios sons eram inadequados ou porque foram usados em locais inadequados, de maneira não sistematizada e sem objetivos explícitos.

Já existem conhecimentos científicos e tecnológicos suficientes para que seja possível iniciar um *pensar sobre o som* como uma das principais dimensões dos ambientes em que habitamos - sejam estes físicos ou virtuais. Isso significa superar a presença do som como ruído⁴ e promover uma atitude orientada ao canal auditivo como portador de informação, apropriando-se dos estímulos sonoros para transmitir mensagens sistematicamente.

6 Referências Bibliográficas

BLATTNER, M.; SUMIKAWA, D.; GREENBERG, R. Earcons and icons: Their

³ Destinado a essa finalidade. Feito exclusivamente para explicar o fenômeno que descreve e que não serve para outros casos, não dando margem a qualquer generalização.

⁴ Segundo [KEEFFE, 2011], o termo ruído é frequentemente usado para descrever sons indesejáveis ou que possuem características que o definem como negativo e desagradável ao ouvinte.

structure and common design principles. **Human Computer Interaction**, 4(1), pp. 11-44, 1989.

BREWSTER, Stephen Anthony. **Providing a Structured Method for Integrating Non-Speech Audio into Human-Computer Interfaces.**

Submitted for the degree of Doctor of Philosophy. University of York, Human-Computer Interaction Group, Department of Computer Science. August, 1994.

CANCELLARO, Joseph. **Exploring Sound Design for Interactive Media.** Delmar Cengage Learning; 1 ed., 2006. Clifton Park, NY.

CARVALHO, Luiz Roberto; PEREIRA, Alice Theresinha Cybis. **O Som Dinâmico em Interfaces Hipermídia: Áudio Interativo Direto e Áudio Adaptativo Indireto.** In: IDEMi Integração para a Inovação - Artigos Selecionados: II Conferência Internacional de Integração do Design, Engenharia e Gestão para Inovação. Albertina Medeiros; Marcelo Gitirana. (Org.). 01ed. Florianópolis: Fundação Universidade do Estado de Santa Catarina, v. 01, p. 246-261, 2012.

COLLINS, Karen. **Game Sound: an introduction to the history, theory, and practice of video game music and sound design.** Massachusetts, USA: MIT Press, 2008.

COLLINS, Karen. Making Gamers Cry: Mirror Neurons and Embodied Interaction with Game Sound. **AudioMostly 2011**, September 7–9, 2011, Coimbra, Portugal.

FOLLETT, Jonathan. **Audio and the user experience.** Disponível em: <http://www.uxmatters.com/mt/archives/2007/06/audio-and-the-user-experience.php> June 18, 2007. Acesso em 17 set. 2015.

GAVER, W. Auditory Icons: Using sound in computer interfaces. **Human Computer Interaction**, 2(2), pp. 167-177, 1986.

GAVER, W. Synthesizing auditory icons. In S. Ashlund, K. Mullet, A. Henderson, E. Hollnagel, & T. White (Ed.), **INTERCHI'93, Amsterdam:** ACM Press, Addison-Wesley, pp. 228-235, 1993.

GAVER, W. The SonicFinder: An interface that uses auditory icons. **Human Computer Interaction**, 4(1), pp. 67-94, 1989.

GAVER, W. Using and creating auditory icons. In G. Kramer (Ed.), Auditory Display, sonification, audification and auditory interfaces. **The Proceedings of the First International Conference on Auditory Display**, Santa Fé Institute, Santa Fé: Addison-Wesley, pp. 417-446, 1992.

GAVER, William W. What in the World Do We Hear? An Ecological Approach to Auditory Event Perception. **Ecological Psychology** 5/1: 1-29, 1993b.

HUG, D. New wine in new skins: Sketching the future of game sound design. In: Grimshaw, M. (Ed.), **Game sound technology and player interaction: Concepts and developments.** Hershey, PA: IGI Global, 2011.

ILARI, Beatriz. Música, comportamento social e relações interpessoais. In: **Psicologia em Estudo**. vol.11 no.1 Maringá Jan./Apr. 2006

JØRGENSEN, K. Time for new terminology? Diegetic and non-diegetic sounds in computer games revisited . In: Grimshaw, M. (Ed.), **Game sound technology and player interaction: Concepts and developments.** Hershey, PA: IGI Global, 2011.

KAUSHANSKY, Karen. **Designing With Audio: What Is Sound Good For?** Disponível em: <http://www.smashingmagazine.com/2012/04/designing-with-audio-what-is-sound-good-for/> April 18th, 2012. Acesso em 17 set. 2015.

KAUSHANSKY, Karen. **Guidelines for designing with audio.** Disponível em: <http://www.smashingmagazine.com/2012/09/guidelines-for-designing-with-audio/> September 14th, 2012b. Acesso em 17 set. 2015.

KEEFFE, Linda O'. Sound is Not a Simulation: Methodologies for Examining the Experience of Soundscapes. In: Grimshaw, M. (Ed.), **Game sound technology and player interaction:**

Concepts and developments. Hershey, PA: IGI Global, 2011.

KONDO, Koji (2007). Painting an Interactive Musical Landscape. **The Annual Game Developer's Conference**, San Francisco, March 4–9.

KUTAY, Steve. **The Game Audio Explosion - A Guide to Great Game Sound Part I: Pre-production and Sound Design.** Mar, 2006. Disponível em: <http://ezinearticles.com/?The-Game-Audio-Explosion---A-Guide-to-Great-Game-Sound-Part-I:-Pre-production-and-Sound-Design&id=158383> Acesso em: 25 ago. 2015.

LILJEDAHN, Mats. Sound for Fantasy and Freedom. In: Grimshaw, M. (Ed.), **Game sound technology and player interaction: Concepts and developments.** Hershey, PA: IGI Global, 2011.

MOSES, Laura. **Sound Design.** EventDV. Vol. 23, Issue 1, p37, Jan/Feb2010.

PARKER, J. R.; HEEREMA, John. Audio Interaction in Computer Mediated Games. **International Journal of Computer Games Technology.** vol. 2008, Article ID 178923, 8 pages. Hindawi Publishing Corporation, 2008.

ROCCHESO, Davide et al. Sonic Interaction Design: Sound, Information and Experience. Conference on Human Factors in Computing Systems: **CHI '08 extended abstracts on Human factors in computing systems.** ACM Association for Computing Machinery, New York, NY, USA, 2008.

SANTOS, Eduardo Toledo. O uso de Metáfora Concreta e Manipulação Direta na Interface do Risko. In: **Hiperfídia: desafios da atualidade/** Vânia Ribas Ulbricht, Alice Theresinha Cybis Pereira (orgs.) – Florianópolis: Pandion, 2009.

SERAFIN, Stefania. Sound Design to Enhance Presence in Photorealistic Virtual Reality. **Proceedings of the 2004 International Conference on Auditory Display.** Sidney, Australia, July 6-9, 2004.

SHNEIDERMAN, Ben; PLAISANT, Catherine. **Designing the User Interface: Strategies for Effective Human-Computer Interaction.** 5a. ed., Prentice Hall, mar., 2009.

SONNENSCHN, David. Sound Design: **The Expressive Power of Music, Voice and Sound Effects in Cinema.** Michael Wiese Productions, 2001.

STEVENS, Richard; RAYBOULD, Dave. **The Game Audio Tutorial: A Practical Guide to Sound and Music for Interactive Games.** Elsevier Inc. 2011.

SUMIKAWA, D.A. **Guidelines for the integration of audio cues into computer user interfaces** (Technical Report No. UCRL 53656). Lawrence Livermore National Laboratory, 1985.

SUMIKAWA, D., BLATTNER, M., JOY, K. & GREENBERG, R. **Guidelines for the syntactic design of audio cues in computer interfaces** (Technical Report No. UCRL 92925). Lawrence Livermore National Laboratory, 1986.

SUNDBERG, J. **Acoustic and psychoacoustic aspects of vocal vibrato.** KTH Computer Science and Communication. Department of Speech, Music and Hearing. STL-QPSR, vol. 35, n. 2-3, pg. 045-068. Stockholm, Sweden, 1994.

VAN LEEUWEN, Theo. **Speech, music, sound.** London, UK: Macmillan, 1999.

WICKENS, Chris; HOLLANDS, Justin G.; BANBURY, Simon; PARASURAMAN, Raja. **Engineering psychology and human performance.** Upper Saddle River, NJ: Pearson; 4rd ed., 2012.

WILHELMSSON, U., & WALLÉN, J. A combined model for the structuring of game audio . In: Grimshaw, M. (Ed.), **Game sound technology and player interaction: Concepts and developments.** Hershey, PA: IGI Global, 2011.