



MIXAGEM DE SOM NA INTERFACE

SOUND MIXING IN THE INTERFACE

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, Design de Som, Áudio Dinâmico

No domínio da interface, uma das principais preocupações dos usuários em relação ao som diz respeito ao incômodo gerado por sons indesejáveis. Historicamente, o som na interface tem sido visto como um aborrecimento, e não como uma solução de comunicação. Nesse sentido, deve-se ser criterioso ao definir focos de atenção quando da utilização do som nos ambientes interativos. O número de objetos auditivos presentes na interface deve ser planejado, e a posição de atenção destes elementos deve apresentar uma escala de grandeza perceptível. Nesse contexto, ao introduzir os conceitos de mixagem de som no escopo dos ambientes interativos, o presente estudo objetiva expor como se dá a organização dos elementos sonoros na interface, de maneira que seja possível obter um direcionamento para um ponto de foco ao utilizar sons, fazendo com que os demais estímulos não compitam entre si, e que mensagens sejam extraídas a partir da utilização dos elementos sonoros da interface.

Interface Design, Sound Design, Dynamic Audio

In the interface domain, one of the main users concern regarding the use of sound is the annoyance generated by unwanted sounds. Historically, the sound in the interface has been seen as a hassle, not as a communication solution. In this sense, it is important to be careful when defining attention points by using sounds in interactive environments. The number of auditory objects present in the interface should be planned, and the attention position of these elements should have a perceptible magnitude scale. In this context, when introducing the concepts of sound mixing in the scope of the interactive environments, this study it aims to expose how the organization of the sound elements in the interface occurs, so that it is possible to obtain a focus point when using sounds, so that the other stimuli do not compete with each other, and messages become extracted from the use of sound in the interface.

1 Mixagem

De acordo com [VALLE, 2009], a mixagem, comumente denominada pela abreviação *mix*, trata-se do procedimento de equilíbrio e ajuste de diversas fontes sonoras, com o objetivo de apresentar todos os seus elementos de maneira clara e inteligível. A mixagem lida com a organização dos elementos sonoros de maneira que seja possível obter um direcionamento para um ponto de foco, de modo que os demais sons não compitam entre si. [HENRIQUES, 2007] aponta que é possível fazer uma analogia aos conceitos de foco e visão periférica para a audição através da mixagem. Apesar de possuímos um campo visual extenso, o foco de atenção se concentra sempre em uma região reduzida, sendo o restante denominado de visão periférica, interpretado pelo cérebro com menos relevância. Porém, quando algo nos chama a atenção no campo periférico, desviamos a atenção para aquele novo ponto. Nesse sentido, é possível afirmar que a audição funciona de modo semelhante à visão.

A função da mixagem é valorizar a mensagem sonora através do uso criativo de diferentes instrumentos, vozes e elementos sonoros que interagem entre si, conferindo-lhes certo grau de concordância. Na mixagem são efetuados ajustes de volume, panorama, equalização e aplicação de efeitos para cada fonte sonora, individualmente, de modo que se estabeleça uma relação de hierarquia e direcionamento comunicativo quando da sua junção. Uma mixagem ideal, com informações interessantes ocorrendo dentro de todo o espectro sonoro, deve conduzir e direcionar o ouvinte para os elementos que deseja destacar e conferir maior grau de relevância. Fazendo uma analogia com a matemática e os eixos cartesianos, [HENRIQUES, 2007] aponta para um cenário de mixagem como possuindo três dimensões:

- Eixo X, de Panorama (ou Pan; Panorâmico): termo usado para designar a localização de um som no campo estéreo¹, definindo sua posição

¹ Estéreo, do inglês *stereo*, trata-se de um processo de reprodução sonora em duas vias (canais) separadas, dando ao ouvinte a sensação

horizontal e graus intermediários. Configuram o palco sonoro no estéreo, como os instrumentos e vozes se distribuem nas saídas de áudio. Gravações em mono (um único canal de saída) não possuem panorama;

- Eixo Y, de Espectro: riqueza de frequências. O campo de frequências presentes na mixagem;
- Eixo Z, de Profundidade: refere-se à proporção de distância que cada instrumento ocupa, de acordo com parâmetros como volume, ambiência, reverberações² e ecos;

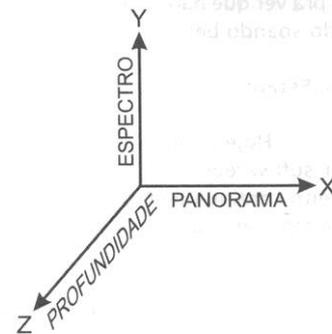


Figura 1: Eixos presentes na mixagem. Fonte: [HENRIQUES, 2007, p. 31].

de posicionamento do som. Quando atinge os ouvidos, o som possui um efeito direcional em relação à esquerda e à direita no espaço, de forma que o som pode estar mais forte de um lado do que de outro. Numa mixagem em estéreo, têm-se duas saídas, esquerda, e direita, e ao usar utilizar estes dois canais é possível criar diferenças de intensidade de um som em cada um dos canais, de forma que o ouvinte será capaz de perceber diferenças de posicionamento e direção das fontes sonoras. A este efeito dá-se o nome de imagem estéreo, que se trata da ilusão do ouvinte ao perceber uma fonte sonora no espaço, quando localizada entre duas caixas acústicas. O termo estéreo é usado genericamente no Brasil, embora a palavra em português seja estereofônico.

² Todos os sons naturais percebidos pelos indivíduos é resultado da soma do som direto com um campo reverberante, composto das reflexões do ambiente. Estas informações psicoacústicas influenciam a percepção de naturalidade, de modo a oferecer uma previsão de tamanho e distância da fonte sonora. A este fenômeno dá-se o nome de reverberação.

[HENRIQUES, 2007] ilustra uma situação em que se assiste ao concerto de uma grande orquestra em um auditório: dezenas de instrumentos estão gerando cada um seu próprio som, somado ao som refletido nas paredes e superfícies do auditório, e os ruídos de fundo, como os sons da plateia. O número de ondas sonoras que chegam aos ouvidos dos espectadores, tanto nesta situação como em qualquer outra, é exatamente o mesmo: apenas uma. A cada instante o ouvido humano recebe uma excitação que é resultante da soma de todas as ondas sonoras que contribuem para seu resultado sonoro. Esta soma não é tão simples, pois envolve grandezas positivas e negativas, o que resulta em reforços e cancelamentos, mas o que o ouvido humano percebe é sempre uma mixagem natural dos sons circundantes, sendo esta mixagem realizada através de dois canais auditivos (ouvido esquerdo e direito), que contribuem para a sensação de posicionamento dos sons no espaço. Por tal razão que, tipicamente, o número de violinos em uma orquestra é maior que o número de oboés, pois as orquestras foram criadas numa época em que não havia meios eletrônicos de amplificação sonora. Quando um instrumento soava mais baixo que o outro, a solução era inserir mais músicos tocando tal instrumento. Os compositores especificavam o número de músicos de cada naipe justamente para prever o volume e a importância daquele instrumento na mixagem da obra, ou seja, no resultado sonoro final. Assim, constata-se que mixar sons sempre foi uma premissa relevante quando do ato de apresentar uma peça sonora coesa e harmônica. Não se trata de um conceito relativamente novo, no entanto, recente foi o desenvolvimento de ferramentas tecnológicas que possibilitaram a execução de mixagens no meio digital, através da utilização de *DAWs*³.

[JORGE, 2002b] aponta que nosso ouvido é capaz de distinguir sons individuais dentro de um conjunto de sons, mas que tal capacidade é

limitada. Só até certo ponto torna-se possível individualizar elementos no contexto de um todo sonoro. Tal capacidade não é considerada muito extensa, tendo em vista que os indivíduos facilmente dispersam sua concentração quando colocados numa situação de conteúdo auditivo excessivo. O número de camadas que o ouvido é capaz de compreender até que surja a incapacidade de distinção é de cinco sons. Isto significa que ao escutar cinco coisas diferentes, o ouvindo humano ainda é capaz de percebê-las individualmente. Quando esta quantidade é ultrapassada, o ouvido humano deixa de compreender cada uma delas, e tem-se, na prática, uma situação de ruído: ouve-se, porém não é possível retirar qualquer informação clara e distinta do que é ouvido. Ao ouvir dez camadas distintas, o indivíduo não será capaz de compreender sequer cinco. Este mesmo som será incompreensível da mesma forma que seis ou vinte camadas emitidas simultaneamente.

Para que esta diferenciação auditiva tenha resultados positivos, [JORGE, 2002b] aponta ser necessário que cada um dos cinco sons ocupe um lugar distinto no espectro auditivo humano. Se assumirmos que o ouvido humano capta sons que se situam entre 20 e 20.000Hz, é necessário que tais sons distribuam-se ao longo de todo o espectro de frequências, pois se todas as camadas situarem-se em torno de uma única frequência, será evidente a dificuldade de compreensão. Além de tais camadas ocuparem posições bem definidas no espectro auditivo, é recomendável que estas possuam uma variação dinâmica de volumes e timbres, a ponto de melhor clarificar cada uma delas. Em suma, tal regra das cinco camadas só ocorrerá se os sons forem efetivamente distintos entre si, como por exemplo, um diálogo, som de chuva e vento, música, motores e buzinas de carros passando na rua e sons de pessoas ao fundo. Tais sons são substancialmente diferentes entre si e cobrem uma margem considerável do espectro de frequências, sendo também sons que não suscitam confusão quanto à sua natureza.

No que diz respeito à audição de sons semelhantes em simultâneo, mantendo sua perceptibilidade, trata-se de uma capacidade relativamente reduzida. [MURCH, 2005] afirma ser possível ouvir dois

³ DAW (do inglês *Digital Audio Workstation*): refere-se aos pacotes de *software* que funcionam como estúdios virtuais de produção musical, compostos de um sistema de gravação e edição de áudio multipista que podem ser operados digitalmente.

sons semelhantes ao mesmo tempo e identificá-los, denominando tal fenômeno como a lei dos dois-e-meio (*two-and-a-half law*; tradução livre). Por exemplo, numa situação em que há vários cães, se um estiver latindo, pode-se distingui-lo perfeitamente, se outro começa a latir, ainda consegue-se identificar cada um deles, porém, se um terceiro cão começar a latir, deixa-se de percebê-los individualmente e tem-se apenas uma ideia de conjunto: há cães latindo. O mesmo vale para os diálogos e para todos os outros sons. Embora com alguma dificuldade seja possível apreender dois diálogos em simultâneo, se uma terceira pessoa começar a falar assume-se esse conjunto de sons como pertencente a um grupo de pessoas falando, onde já não é possível distinguir seu conteúdo.

1.1 Equalização

De acordo com Henriques (2007, p.67), “equalizar significa igualar, tornar igual, sem variações”. A equalização diz respeito à alteração nos valores das frequências (graves, médios e agudos) de determinando som, de modo a destacar algumas regiões de frequências ou evitar o congestionamento das mesmas no espectro sonoro. Nesse sentido, a equalização pode ser do tipo corretiva, objetivando reduzir problemas advindos de imperfeições, como ruídos ou o excesso de determinadas frequências, ou criativa, visando enaltecer regiões do espectro para obter resultados estéticos específicos. Leo de Gars Gulka (*apud* Henriques, 2007, p.63), engenheiro de áudio que se destacou na década de 1970, dividiu o espectro de frequências percebidas pelo ouvido humano nas seguintes regiões:

20 a 60 Hz	Sub-Graves
60 a 250 Hz	Graves
250 a 2000 Hz	Médias Baixas
2000 a 6000 Hz	Médias Altas
6000 a 20000Hz	Agudos

Figura 2: Espectro de frequências segundo Leo de Gars Gulka. Fonte: [HENRIQUES, 2007, p. 63].

[HENRIQUES, 2007] assinala que até meados dos

anos 90, a região de sub-graves era mais um inconveniente do que um benefício numa mixagem. Trata-se, pois, de uma região difícil de conseguir nos sulcos de um LP de vinil, mas com o advento de novas mídias e formatos de amplificação capazes de responder a essa região com alguma eficiência, tornou-se possível atuar com propriedade na zona dos sub-graves. São frequências muito mais sentidas pelo seu poder de vibração do que propriamente percebidas ao ouvido. Segundo o autor, os sub-graves são cativantes pois nos remetem ao útero, onde todos os agudos do mundo externo são filtrados e ouve-se constantemente o pulsar do coração materno. A indústria cinematográfica se aproveitou desta característica e criou um canal de áudio para as mixagens de cinema especialmente dedicado a gerar efeitos com alto poder de vibração, chamado *LFE* (*low frequency effects* – efeitos de baixa frequência), usados para conferir mais intensidade a momentos como explosões, sons de naves espaciais, navios, terremotos. Demais aparelhos de amplificação de som herdaram este canal, sendo o mesmo inserido na saída de áudio responsável também pela geração das demais frequências graves, sendo conhecido como *subwoofer*⁴. Pelo fato das frequências graves gerarem um impacto de presença e força sonora nas composições, a demanda popular pela sua audição - inclusive em mixagens sem uma notória adição de graves - ocasionou o surgimento de uma função que adiciona frequências graves nos aparelhos de som, possuindo nomenclaturas distintas⁵ de acordo com cada fabricante.

A maior parte da energia dos instrumentos e vozes utilizados para compor a base das músicas se encontra na região de frequências médias-baixas. Deve-se ter cautela nessa zona, pois seu excesso pode ser percebido pelo ouvido como falta de

⁴ Tipo de caixa acústica projetada para reproduzir frequências extremamente baixas, de 20Hz a 100Hz. Essas frequências em geral são mais sentidas do que propriamente ouvidas; as caixas de *subwoofers* podem ficar localizadas praticamente em qualquer lugar do ambiente, pois o ouvido não percebe claramente a origem de sons graves.

⁵ *Superbass, extrabass, extremebass, powerbass, ultrabass*, entre outros.

frequências médias-altas e agudos, dando a impressão de que o som está *velado, coberto*. Entretanto, quando um instrumento soa sem *peso*, é esta a região que deve ser acentuada. Por tratar-se de uma região do espectro de áudio que coincide com boa parte das frequências emitidas pela voz humana, para que a mesma soe inteligível e se destaque no contexto da mixagem, recomenda-se a atenuação dessas frequências por parte de outros sons que estejam competindo com a voz principal da mixagem - mesmo que a *voz principal* não seja necessariamente a voz humana. Porém, deve-se ter cautela para que a mixagem não soe sem *peso* por conta de atenuações exageradas nessa região do espectro de frequências.

As médias-altas formam uma região extremamente importante em uma mixagem. Primeiro, porque o pico de sensibilidade do ouvido está nessa região - entre 2 e 3 kHz, aproximadamente -, e também porque na sua metade superior - entre 4 e 6 kHz -, se encontra uma região que é denominada de *presença*, na qual os detalhes sonoros são mais bem percebidos. Por atuar no pico da sensibilidade auditiva, acentuar ou reduzir o ganho na região entre 2 e 4 kHz deve ser um processo conduzido com cuidado, pois o resultado pode facilmente tornar-se desagradável. É nesta região onde as consoantes se tornam nítidas (o que diferencia um F de um S são justamente os harmônicos desta região), sendo uma ótima região para tornar um elemento sonoro mais definido e com mais destaque na mixagem.

[HENRIQUES, 2007] aponta que o termo técnico denominado “*to cut through the mix*” é utilizado para definir o ato de tornar um elemento sonoro mais aparente no meio dos outros - e trabalhar com acentuações de ganho na região entre 2 e 4 kHz é uma opção efetiva de fazê-lo. Os sons que povoam a região dos agudos são normalmente de curta duração, tais como a consoante S da voz, pratos de bateria e demais harmônicos altos de outros instrumentos. O excesso de informação nestas frequências pode levar à fadiga auditiva, tornando a audição da composição sonora desconfortável. A região superior de agudos - acima de 15 kHz - é comumente chamada de *ar*.

[TONG e WONG, 2006] apontam que é possível usar parâmetros de equalização para que um som se torne mais evidente, como ocorre ao acentuarem-se as frequências que se encontram em torno de 3.000Hz num diálogo. Como resultado, tem-se um som mais claro e audível. A velocidade de reprodução de um som pode também alterar sua propriedade. Para um som com características rítmicas, como o som emitido por um motor de automóvel, a alteração da sua velocidade gera uma interpretação acerca da quantidade de energia produzida. A redução de seu tempo irá resultar numa sensação de perda de energia, enquanto que o aumento de sua velocidade é um indicativo de potência e vigor.

2 Mixagem de Som na Interface

Uma das principais preocupações dos usuários em relação ao som na interface diz respeito ao incômodo gerado por sons indesejáveis, como apontado por [BREWSTER, 1994] desde o início dos estudos do som na IHC. Desde então, o som na interface tem sido visto como um aborrecimento, não como uma solução de comunicação. Dada à natureza intrusiva do áudio, [FOLLETT, 2007] pontua que é importante pedir permissão ao usuário para utilizar estímulos sonoros, e sempre manter a possibilidade do som ser completamente desligado, a qualquer instante. Nesse contexto, há dois aspectos a se considerar: os sons podem ser irritantes para o usuário, ou para os demais indivíduos no mesmo ambiente em que os sons são disparados. Deve-se, portanto, ser criterioso ao definir focos de atenção, pois o som pode se tornar desagradável e precisa ser utilizado com cuidado para transmitir o efeito desejado.

Em ambientes interativos, deve-se utilizar o som para criar pontos de foco de tal modo que, ao invés de reunir uma cacofonia de ruídos concorrentes, obtenha-se um direcionamento sonoro que atue em conjunto à interface para atingir uma proposta comunicacional definida. O número de objetos auditivos presentes na interface deve ser planejado cuidadosamente, tendo em vista que seu excesso pode causar um impacto negativo, resultando no que é denominado por *poluição sonora*. Nesse sentido, um equilíbrio constante entre todos os

elementos sonoros da interface não irá direcionar nenhum tipo de foco para um aspecto específico da interface, no entanto, se determinado som muda abruptamente de volume, atrai a atenção dos usuários, e torna-se mais importante. Utilizando-se o volume⁶ podem-se classificar quais sons são importantes e quais possuem menor relevância - trata-se da dinâmica de controle de volume que é mencionada em Gibson (1997).

De acordo com [SERAFIN, 2004], quando se trata da utilização de sons em sistemas interativos, muitos conceitos de design de som podem ser utilizados da indústria do entretenimento, que há muito já reconheceu a importância do som para adicionar presença e imersão no cinema e nos games. Nos filmes, os efeitos sonoros exageram a realidade para criar uma experiência imersiva – do mesmo modo, ambientes virtuais são criados para que os usuários possam experimentar uma complexa experiência de imersão. Nesse sentido, [MOSES, 2010] aponta que o significado de uma mensagem pode ser transmitido através do emprego de sons hiper-reais. O som exagerado pode ser útil para transmitir uma mensagem, pois ao elevar o nível do som além do que se poderia experimentar na realidade, sua magnitude e importância são enfatizadas. Isto se dá porque o cérebro humano, naturalmente, entra em sintonia com os sons que são pertinentes e coloca em segundo plano o restante, e ao exagerar alguns sons, impede-se que os mesmos sejam postos em segundo plano.

Sobre a dinâmica existente entre som e imagem no contexto da narrativa fílmica, [JORGE, 2002] aponta duas noções - pontuação e amplificação - que são pertinentes para um melhor esclarecimento da relação do som na hipermídia. Por *pontuação* entende-se a capacidade que o som tem de acentuar

⁶ Intensidade do som, que está diretamente associada ao deslocamento de ar provocado pela onda sonora. O ouvido humano pode perceber sons dentro de uma faixa muito grande de intensidades, que vão desde o limiar da audição (0 dB SPL) até o estrondo de um avião a jato (130 dB SPL). A qualidade da sensibilidade do ouvido varia de acordo com o volume do som: para volumes muito baixos, a sensibilidade é relativamente menor nas frequências extremas (graves e agudos).

um objetivo, uma ideia a ser transmitida. Isto consiste em tornar identificável, através de sons, qual o sentido comunicativo pretendido na interface. Se o objetivo é transmitir alegria, por exemplo, os elementos sonoros devem ser direcionados para tal intenção. Por *amplificação* entende-se a capacidade que o som tem de transportar os indivíduos além de imagens, ou até mesmo, de alterar o próprio sentido da imagem. Constata-se o uso da amplificação sempre que sons são usados para atingir um sentido que os transcenda, podendo chegar ao ponto dos próprios sons significarem qualquer coisa que, manifestamente, não possuem relação com a imagem.

[DROUMEVA, 2011, p. 139] elenca posições de escuta de acordo com a intensidade destes sons, ordenando três planos de atenção (primeiro, segundo e plano de fundo) de acordo com as funções comumente atribuídas a estas posições nos sons utilizados em games. A figura a seguir sumariza tais apontamentos:

Posição de atenção	Função na Interface	Tipo de trilha sonora
1 PRIMEIRO PLANO	Funções direcionadas à ação.	Alertas: notificações de confirmação ou rejeição de ação e de status de interface.
2 SEGUNDO PLANO	Funções de orientação e identificação.	Efeitos sonoros de contexto: auditory icons e earcons.
3 TERCEIRO PLANO	Funções de atmosfera.	Música e ambiência.

Figura 3: Posições de escuta de acordo com a intensidade de sons na interface. Fonte: adaptado de [DROUMEVA, 2011, p. 139].

[WILHELMSSON e WALLÉN, 2011] pontuam que efeitos sonoros em primeiro plano são mais eficazes em evocar emoções súbitas no usuário do que sons alocados no segundo plano, ou no plano de fundo. Sons mais suaves, no entanto, podem atuar como ferramenta atmosférica, sendo capazes de aumentar a imersão e definir um estado de humor para os usuários. [TOPRAC e ABDEL-MEGUID, 2011] assinalam que games geralmente utilizam a ambiência como sons de fundo, e efeitos



sonoros para evocar emoção em volumes médios e altos, de modo a maximizar sua percepção pelo usuário.

Em games, bem como nos filmes, os sons em si são em grande parte construídos e montados. [COLLINS, 2008] observa que particularmente nos games, os sons dificilmente são sons reais gravados, mas geralmente uma construção simulada de sons digitais sintetizados que remetem a um simulacro do real. Até mesmo os próprios sons reais gravados, tais como sons de explosões, são muitas vezes tratados com vários efeitos e filtros para torná-los “mais reais do que o próprio som real”. No entanto, existe uma diferença fundamental entre as aplicações de som nos filmes e nos sistemas interativos. Enquanto designers de som de filmes podem criar uma trilha sonora de duração limitada e correspondente aos elementos visuais, nos ambientes interativos as características de comportamento dos usuários podem variar substancialmente. Ou seja, nos sistemas interativos, um campo sonoro dinâmico deve ser criado de tal modo que possa responder adequadamente às mudanças de rota de navegação provocadas pelo usuário.

2.1 Silêncio

O som em ambientes interativos é descrito como pertencente a três elementos: diálogo, ambiência e efeitos sonoros. [Liljedahl, 2011] argumenta que a ausência de som pode resultar numa representação inadequada e incompleta de um ambiente virtual, dada à inexistência do senso de presença que é percebido pelos usuários através do uso de sons. No entanto, [MCKEE, 2006] aponta que o silêncio, ou a quase ausência de som, precisa também fazer parte de quaisquer considerações sobre som. O silêncio era, até recentemente, a configuração padrão para qualquer exibição interativa na *web*, tendo em vista que sua reprodução era muito fragmentada e os métodos para inserir arquivos de som muito complexos. Porém, tais condições têm mudando nos últimos anos e o silêncio já não é necessariamente a configuração padrão de exibições interativas.

Sendo possível ouvir sons, pode-se também *ouvir o*

silêncio, ou seja, sua utilização torna-se propositada. O silêncio não é mais um padrão, mas uma escolha, e designers de sistemas interativos necessitam utilizá-lo conscientemente. Perceber o silêncio é algo complexo, pois envolve a escuta de uma presença ausente - o silêncio legitima o seu rompimento e o som legitima o silêncio em tal intercalação que não há, de fato, uma separação entre eles - daí a importância do silêncio em todas as discussões sobre som. O silêncio não deve ser considerado separadamente, mas como importante integrante da trilha sonora da interface, cuja relação com os demais elementos diálogo, ambiência e efeitos sonoros, deve ser atendida. O som pode ser estrategicamente utilizado para enriquecer a estrutura geral de uma interface, no entanto, deve-se considerar que a ausência de som é também capaz de transmitir informação ao usuário, como por exemplo, indicar que este já concluiu alguma tarefa específica da interface, e deve sair de determinada área, progredindo pelos links sugeridos. Do mesmo modo, uma pausa pode indicar uma mudança na narrativa da interface, bem como o uso de música contínua em interfaces díspares pode ajudar a sinalizar na continuação de um tema específico de conteúdo.

É impossível para os indivíduos com capacidade auditiva normal experimentar o silêncio total. Não existe silêncio absoluto, ou a total ausência de estímulos auditivos. [BLESSER e SALTER, 2007] pontuam que numa câmara anecóica⁷ nossas funções corporais internas tornam-se audíveis, como o som do coração, do fluxo sanguíneo, do sistema digestório. A percepção auditiva é subjetiva, e o silêncio é sempre relativo a algum outro som que é mais acentuado que o próprio silêncio percebido. Os seres humanos não possuem um nível fixo de referência para sons altos e baixos, pois percebem o som apenas em nível comparativo, nunca absoluto. Por exemplo, quando um indivíduo está em silêncio no meio da noite e acidentalmente derruba um copo de vidro, este som

⁷ Ambiente construído artificialmente para apresentar reflexão sonora nula, e isolamento acústico total. Trata-se de uma cápsula que não possui nenhum tipo de troca vibracional com o ambiente externo, nem tampouco reflete ondas sonoras em seu interior.

é considerado extremamente alto. Ao caminhar rumo a uma aeronave para o embarque, o indivíduo percebe o som da turbina como alto, mas é incapaz de distinguir claramente as diferenças de pressão sonora exercida por ambos os estímulos, e apenas pontua que ambos são estímulos intensos. Em termos de nível de pressão sonora (do inglês *sound pressure level*), o motor a jato é mais intenso que o som do copo quebrando, no entanto, no momento em que os eventos ocorreram, seus referenciais de som e da ausência de som fizeram com que o som do copo fosse percebido com tanta intensidade quanto o som da turbina da aeronave.

3 Considerações Finais

Possibilidades significativas de melhorias no campo do design de som são dadas quando o som é analisado sob uma ótica abrangente no que diz respeito às suas características de conteúdo, forma e função. Em termos de percepção, tais parâmetros sonoros são difíceis de estimar com alta precisão, tornando-se ainda mais complexos quando no domínio da hiperfídia, que dão origem a uma equivalente perceptual que somente corresponde a uma realidade factual física aproximada, tendo em vista que muitos sons presentes nos dispositivos de IHC não possuem uma correspondente física, pois foram criados exclusivamente para utilização na interface. A questão central, portanto, deve concentrar-se em apontar como os ouvintes processam os eventos acústicos da interface quando estes são tidos como processos de comunicação e, portanto, portadores de informações específicas.

Ao pontuar a mixagem 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. 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*⁸ 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.

A duração de determinado estado de interface é um elemento complexo de se quantificar, tendo em vista que cada usuário a manipula de acordo com seu interesse. Logo, a experiência do usuário é fator determinante para que haja interatividade com os elementos sonoros presentes na interface. Para que o design de som venha a se integrar, de fato, ao design de interface, a solução mais apropriada é oferecer aos usuários o total controle sobre as possibilidades de som, tornando a abordagem do áudio dinâmico consonante com os princípios do design para a experiência do usuário. 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 ruído⁹ e promover uma atitude orientada ao canal auditivo como portador de informação, apropriando-se dos estímulos sonoros para transmitir mensagens sistematicamente.

4 Referências Bibliográficas

ALZAMORA, Geane Carvalho. Fluxos de informação no ciberespaço – conexões emergentes.

⁸ 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.

⁹ De acordo com 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.

Revista Galáxia, São Paulo, n. 13, p. 75-88, jun. 2007.

BLESSER, B; SALTER, L. **Spaces speak, are you listening? Experiencing aural architecture**. The MIT Press; Cambridge, MA, 2007.

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.

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

DROUMEVA, Milena. An Acoustic Communication Framework for Game Sound: Fidelity, Verisimilitude, Ecology. In: Grimshaw, M. (Ed.), **Game sound technology and player interaction: Concepts and developments**. Hershey, PA: IGI Global, 2011.

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.

HENRIQUES, Fábio. **Guia de Mixagem**. 2ª Ed. Rio de Janeiro: Música e Tecnologia, 2007.

JORGE, Rui Pereira. A música dos videogames. **Caleidoscópio: Revista de Comunicação e Cultura**. Lisboa, Portugal: Departamento de Ciências da Comunicação, Artes e Tecnologias da Informação da Universidade Lusófona de Humanidades e Tecnologias (ULHT). Vol.2, n.2, 2002.

JORGE, Rui Pereira. Edição de som: algumas perspectivas. **Caleidoscópio: Revista de Comunicação e Cultura**. Lisboa, Portugal: Departamento de Ciências da Comunicação, Artes e Tecnologias da Informação da Universidade Lusófona de Humanidades e Tecnologias (ULHT).

Vol.2, n.2, 2002b

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

MCKEE, Heidi. Sound matters: Notes toward the analysis and design of sound in multimodal webtexts. **Computers and Composition**. Vol. 23, Issue 3. p. 335-354. [s.l.]: Elsevier / ScienceDirect, 2006.

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

MURCH, Walter. **Dense Clarity – Clear Density**. The transom review, Part 2, Vol. 5/Issue 1, April 1st, 2005. Disponível em: http://transom.org/?page_id=7006. Acesso em: 05 set. 2015.

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.

TONG, Kam-pang Maya; WONG, Kam-wah. Schematic interface of sound creation for computer animators. **Journal of Zhejiang University**. China, Science A. vol. 7, n. 7, 2006.

TOPRAC, Paul; ABDEL-MEGUID, Ahmed. Causing Fear, Suspense, and Anxiety Using Sound Design in Computer Games. In: Grimshaw, M. (Ed.), **Game sound technology and player interaction: Concepts and developments**. Hershey, PA: IGI Global, 2011.

VALLE, Sólton do. **Manual Prático de Acústica**. 3ª Ed. Rio de Janeiro: Música e Tecnologia, 2009.

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.