

A INTERAÇÃO SONORA COM SISTEMAS EMBARCADOS EM VEÍCULOS

SONIC INTERACTION WITH IN-VEHICLE SYSTEMS

Manuela Quaresma¹, D.Sc.
Isabela Motta², graduanda em Design

(1) LEUI / Laboratório de Ergodesign e Usabilidade de Interfaces – PUC-Rio
e-mail: mquaresma@puc-rio.br

(2) LEUI / Laboratório de Ergodesign e Usabilidade de Interfaces – PUC-Rio
e-mail: isabela.canellas@gmail.com

Interação Sonora, Sistemas Embarcados em Veículos, Distração do Motorista

Com o aumento do uso de sistemas embarcados para motoristas e seu potencial de distração, é necessário buscar outras abordagens que não ocupem o canal visual, uma vez que ele deve estar voltado para a condução. Este artigo apresenta uma revisão de literatura sobre a interação do motorista com interfaces sonoras, a fim de investigar os impactos dessa interação no contexto da condução e ressaltar algumas diretrizes de design para essas interfaces. Foram identificadas cinco categorias que combinam interfaces para motoristas e diferentes tipos de interação sonora, resultando na identificação e discussão de nove diretrizes sobre o assunto.

Sonic Interaction, In-Vehicle Systems, Driver Distraction

With the increased use of in-vehicle systems for drivers and their potential for distractions, it is necessary to seek other ways that do not involve the driver's visual channel, since it should be dedicated to driving. This paper presents a literature review on the interaction of the driver with sonic interfaces to investigate the impacts of this interaction in the driving context and to highlight some design guidelines for these interfaces.. Five categories were identified that combine interfaces for drivers and different types of sonic interaction, resulting in the identification and discussion of nine guidelines on the subject.

1 Introdução

O uso de tecnologias embarcadas e dispositivos móveis em veículos vem crescendo rapidamente nos últimos anos. De acordo com a Organização Mundial da Saúde, em 2011, proporção de motoristas que utilizavam seus telefones celulares enquanto dirigiam tinha aumentado consideravelmente em 5-10 anos, num percentual de 1% para 11% (OMS,2011). Porém, o

desempenho de tarefas paralelas à condução do veículo é um grande risco para o motorista, pois podem o distrair na principal tarefa de dirigir.

Yang et al. (2012) observaram que tarefas secundárias aumentam a demanda pelos recursos cerebrais do motorista. Segundo Horrey e Wickens, (2006) apud Kun, Shyrovkov e Heeman (2010), chamadas telefônicas, mesmo em modo *hands-free*, podem prejudicar o desempenho do motorista.

Pode-se argumentar que um dos motivos para a piora na direção quando há uma tarefa secundária acontecendo é devido à alocação da atenção visual do motorista da via para os dispositivos dentro do veículo. Kun et al. (2009) observaram que a piora do desempenho do motorista estava sempre associada ao tempo que o motorista desviou o olhar da via.

Mesmo sendo uma tarefa predominantemente visual, outros tipos de distração também podem afetar o desempenho da direção. De acordo com a NHTSA – National Highway Traffic Safety Administration (NHTSA, 2012), órgão responsável pela segurança no trânsito norte-americano, a distração do motorista pode ser definida como “um tipo específico de desatenção que ocorre quando os motoristas desviam sua atenção da tarefa de dirigir para se concentrar em outra atividade”. Esta distração pode ocorrer de diferentes maneiras, de acordo com as seguintes categorias: *visual* – através de tarefas que exigem que o motorista desvie o olhar da via para obter informação visual; *manual* – através de tarefas que exigem que o motorista retire uma das mãos do volante para manipular algum dispositivo; e *cognitiva* – através de tarefas que tiram a atenção do motorista para longe da tarefa de dirigir.

Segundo a Teoria de Múltiplos Recursos (*Multiple Resource Theory*), áreas diferentes do cérebro são responsáveis por diferentes e limitados recursos de processamento, códigos e modalidades. Quando múltiplas tarefas utilizam recursos diferentes, não competindo pelas mesmas áreas do cérebro, o compartilhamento de tempo e atenção entre elas será mais eficiente e menos comprometedor. Por outro lado, quando duas tarefas usam o mesmo recurso elas tendem a competir entre si, prejudicando o desempenho do motorista (Wickens, 2008).

Wickens (2008) sugere que a utilização de tais recursos, por serem limitados, está ligada a carga mental sofrida pelo usuário. Caso a demanda das tarefas seja menor do que os recursos disponíveis, uma capacidade residual sobrar para situações inesperadas, as quais podem ocorrer ao dirigir. Porém, caso exceda os recursos disponíveis, o

desempenho será afetado.

Tendo em vista o risco de a distração do motorista ser primordialmente visual, pode-se pressupor que o uso do som como forma de interação para motoristas é uma boa opção, na medida que não compete pelas mesmas modalidades cognitivas do condutor, atenuando a sobrecarga no canal visual. Segundo Brewster (1997) e Brown, Newsome, Glinert (1989) apud Jeon et al. (2009), a apresentação da informação pelo canal auditivo facilita o desempenho com interfaces, podendo evitar uma sobrecarga visual. Além disso, comparada à percepção visual, a audição tem características vantajosas como a localização espacial, a intuitividade e a rapidez na percepção e reação (Wang, Li e Chen, 2012).

Entretanto, ainda há necessidade de investigar os impactos gerados pelo uso do som em interfaces na condução. Um som pode dominar ou causar um breve lapso no desempenho da tarefa visual, uma vez que a modalidade auditiva tem maior capacidade de atrair atenção (Horrey e Wickens, 2004; Wickens e Liu, 1998; Spencer e Driver, 1997; Stanton, Booth e Stammers, 1992 apud Jeon et al., 2009). Porém, ao comparar diversas modalidades visuais com a modalidade auditiva, Cao et al. (2009) observaram que a modalidade visual de ícones foi percebida mais rapidamente que a auditiva, mostrando que a efetividade das modalidades ainda deve ser discutida.

2 A interação através do som

É possível observar que o uso do som pode trazer desvantagens, principalmente quando mal projetado e irritante, mas também é uma ferramenta que pode trazer benefícios para a segurança do motorista. Sobre esse tópico, resultados interessantes foram obtidos por Weinberg, Harsham e Medenica (2011), que, apesar de terem observado um aumento do tempo de realização da tarefa secundária, observaram uma melhora no desempenho da tarefa principal e um menor nível de carga mental para a condição auditiva, quando comparada às duas visuais.

O som está sempre presente no dia-a-dia, e

estamos constantemente produzindo sons resultantes de nossas ações. Essa natural resposta sonora gera uma percepção de causalidade, que é constantemente interpretada pelo cérebro, o qual cria uma imagem de como aquele som foi originado, e ajuda o ser humano a adaptar seus movimentos e gestos conforme o som escutado. Porém, mesmo com seu caráter natural, o som muitas vezes não é conscientemente percebido pelos seres humanos (Hermann et al., 2011).

Por uma definição provinda da física, o som é uma onda, e uma onda sonora é composta de algumas características que podem ser modificadas para projetar interações auditivas em interfaces. A variação da frequência resulta em diferentes *pitches*, que podem ser mais baixos ou altos (graves ou agudos). O volume do som é resultado da variação da amplitude de uma onda. O timbre é a assinatura sonora única de um som, seja ele feito por um instrumento ou pela voz de uma pessoa.

A manipulação das características sonoras e da percepção de causalidade e discriminação de sons pelo ser humano possibilitam a criação e aplicação de uma variedade de sons, tornando-o em uma ferramenta potente para o *design* de interação.

O design de interação sonora (*Sonic Interaction Design*, ou SID) "usa o som para comunicar uma informação, um significado, uma estética ou emoção em contextos interativos. Seu desafio é criar interações adaptáveis que respondam constantemente aos gestos do usuário" (Hermann et al., 2011, tradução nossa). É possível perceber a interação sonora em diversas mídias, como os *feedbacks* sonoros gerados por computadores, micro-ondas, alarmes, jogs etc.

Entretanto, uma consideração importante do design de interação sonora é o contexto no qual a informação é apresentada. É o contexto que irá limitar ou ampliar as possibilidades as quais uma informação pode ser apresentada, dependendo do tipo de interação e do ambiente sonoro no qual o usuário se encontra (Carvalho e Pereira, 2014). No contexto de direção, o barulho do carro pode atrapalhar no reconhecimento da voz do motorista pelo sistema (Tashev, Seltzer, e Ju, 2009; Cao et al., 2010).

3 Os Tipos de Interação Sonora

No design de interação sonora, o som pode ser projetado para ser emitido e escutado por algumas modalidades:

3.1 Comandos de Voz - Entrada e Saída de Voz

Interação sonora por comandos de voz acontecem quando é realizado um comando pelo usuário para o sistema (entrada) ou quando o sistema fornece uma orientação ou uma resposta à uma ação do usuário ou dele mesmo (saída), sendo essa interação de forma oral e verbal. Muito comuns em sistemas *GPS's*, os comandos de voz podem ser apresentados de forma totalmente auditiva ou em conjunto com um aparato visual. Segundo Roginska (2013), os comandos de voz têm um grande nível de especificidade, um alto grau de reconhecimento da mensagem, e precisam de pouco ou nenhum tempo de treinamento para serem aprendidos. No entanto, podem ser lentos e atrapalhar ou distrair o usuário se ele estiver envolvido em uma tarefa de voz ou verbal, como uma conversa entre o motorista e o passageiro, por exemplo. Além disso, para processos contínuos a saída de voz pode ser intrusiva.

3.2 *Spearcons* e *Spindexes*

Uma variação das saídas de voz, *Spearcons* são curtos cliques de áudio gerados automaticamente usando uma síntese de texto-para-voz, de forma a praticamente não serem mais identificáveis como voz (Roginska, 2013). Assim como os comandos de voz em velocidade normal, os *Spearcons* não precisam de muito tempo ou treinamento para serem aprendidos.

Spindexes são curtos sinais de áudio que geralmente representam a primeira letra de uma palavra em um menu. Então, em um menu de contatos, por exemplo, uma letra específica representará todos os contatos que começam com essa letra (Roginska, 2013). Dessa forma, o usuário consegue navegar em grandes listas de forma mais rápida em comparação à saída de voz comum ou aos *Spearcons*. Depois de entrar na zona desejada, o usuário precisa ser apresentado à informação de forma mais detalhada (Roginska, 2013). Esse

detalhamento na exposição dos dados pode ser feito por voz, ou até mesmo com *Spearcons* (Jeon et al. 2009).

Por serem conceitos recentes, esse tipo de interação sonora ainda não foi muito explorado, e é utilizado principalmente em menus e listas, os quais se beneficiam da velocidade da interação em relação aos comandos de voz comuns.

3.3 Ícones Sonoros e Ícones auditivos

Ícones sonoros (*earcons*) são mensagens de áudio geradas artificialmente, não-verbais e abstratas, geralmente de duas a quatro notas, usadas para apresentar informação sobre um objeto, uma operação ou uma interação (Roginska, 2013). Por serem abstratos e de difícil interpretação, os ícones sonoros são usados para prover informações curtas e simples atraindo a atenção do usuário para a ação, mas, ao contrário dos comandos de voz, têm baixo grau de especificidade. Esse tipo de interação pode ser usado em monitoramentos de uma ação ou processo, como os utilizados em hospitais e em alarmes de colisão de automóveis, mantendo o usuário constantemente informado de uma ação que está ocorrendo, se adaptando às suas ações e apresentando informações atualizadas.

Ícones auditivos (*auditory icons*) são os ícones sonoros (*eacons*) representativos, sons do dia-a-dia que são retirados do seu contexto original e atribuídos a uma nova ação computacional com um novo significado (Roginska, 2013). A associação do som de um papel sendo amassado a um arquivo digital sendo deletado é um exemplo de ícone auditivo. Logo, os ícones auditivos são intuitivos e precisam de menos tempo de aprendizado, uma vez que já estão originalmente relacionados a uma ação cotidiana (Roginska, 2013).

4 Metodologia

O objetivo desse artigo é identificar o que está sendo discutido sobre a interação com interfaces sonoras em automóveis para 1) investigar quais são os impactos, sejam eles positivos ou negativos, dessa interação no contexto da condução e 2) evidenciar diretrizes para o design de interação sonora para motoristas.

Com esse objetivo, foi feita uma revisão de literatura que analisou publicações dos anais da *AutomotiveUI – International Conference on Automotive User Interfaces and Interactive Vehicle Application*, dos anos de 2009 a 2015. A *AutomotiveUI* é uma conferência internacional sobre interfaces para motoristas, e por isso foi escolhida como delimitadora de escopo para este artigo. As publicações foram escolhidas de acordo com as palavras chave em seus resumos. As palavras foram: *Speech, Spoken, Voice, Earcons, Auditory, Audio, Sound, Spearcon, Spindex*. Foram recolhidas 37 publicações, entre artigos e resumos de pôsteres, sendo duas descartadas por inadequação ao escopo de pesquisa, totalizando 35 publicações analisadas.

As publicações foram analisadas, primeiramente, em relação a adequação ao tema e ao escopo da pesquisa. Depois disso, foram identificadas as palavras e conceitos chaves em relação aos tipos de interação sonora (Comandos de Voz, *Spindex e Spearcons*, Ícones sonoros e Ícones auditivos) e aos tipos de interfaces para motoristas, os ADAS (*Advanced Driver Assistance System*) ou sistemas avançados de assistência ao motorista, e os IVIS (*In-Vehicle Information Systems*) ou sistemas de informação embarcados.

Os ADAS são “sistemas que interagem com o motorista com o propósito de dar suporte à tarefa de condução, em níveis de monitoramento e regulação” (Östlund et al., 2005, tradução nossa). Exemplos de ADAS são: avisos de colisão frontal ou traseira, avisos de desvio na faixa, etc.

Por sua vez, IVIS são “sistemas que interagem com o motorista com o propósito de dar suporte a tarefas de identificação de alvos e de monitoramento, ou que não dão suporte nenhum à tarefa de condução” (Östlund et al., 2005, tradução nossa). Exemplos de IVIS são: sistemas de monitoramento de gasolina e de pressão dos pneus, GPSs, monitoramento do funcionamento do veículo, rádios, notícias, emails, mensagens etc.

Após a identificação dos conceitos chave, o conteúdo dos artigos foi categorizado segundo o cruzamento dos tipos interações sonoras com os tipos de interface para motoristas, quando havia

relação possível entre elas. Essa divisão resultou em um total de cinco categorias, as quais serão analisadas na próxima seção.

5 Resultados

Ao observar o conteúdo revisado, foram identificados parâmetros com os quais a interação sonora é construída e analisada pelos pesquisadores, sendo os mais frequentes utilizados como objeto de estudo deste artigo. Os parâmetros a seguir foram utilizados na análise das categorias com o objetivo de identificar diretrizes sobre a interação sonora para motoristas.

Desempenho da Tarefa Primária – desempenho do motorista na tarefa principal de dirigir, medidos a partir de: tempo e número de olhadelas para dentro do carro; desvio da posição na faixa e ângulo do volante; número de colisões ou “quase-colisões”; número de erros de uma tarefa específica; tempo de reação; e velocidade média e distância do veículo à frente. Os experimentos observados nas publicações usaram testes nas vias, simuladores de alta, média e baixa fidelidade, testes como o LCT - *Lane Change Test* (Teste de Troca de Faixa), jogos, desafios de atenção e outros para engajar os participantes em uma tarefa equivalente à condução.

Desempenho da Tarefa Secundária – desempenho do motorista em qualquer tarefa que não seja a direção, como: conversas telefônicas; manuseio ou uso por voz de celulares; utilização de GPSs; sintonização de rádio; e navegação em listas por comandos de voz ou manualmente.

Carga Mental – medida pelos autores usando questionários como o NASA-TLX, DALI, RSME.

Interpretação Subjetiva – como a percepção de urgência; percepção de irritabilidade; percepção de efetividade; taxa de reconhecimento de uma mensagem; e preferência

Multimodalidade – revisando o conteúdo das publicações, foi observado que em alguns experimentos foram comparadas mais de uma modalidade de interação (ex: visual, auditiva, tátil)

de forma isolada ou mesclada, o que será tratado neste artigo como Multimodalidade.

Flexibilidade de aplicação – conceito de flexibilidade de aplicação será discutido de forma a responder em que momentos um tipo de interação pode ser usado da melhor maneira, e de que forma é possível variar o uso de um ou mais tipos de interação sonora em um ambiente de condução. Esse conceito está relacionado ao uso de interfaces adaptativas.

As categorias identificadas serão apresentadas a seguir, expondo, primeiramente, as diretrizes observadas, seguidas da discussão que levou a essas conclusões.

5.1 Earcons e Ícones auditivos em ADAS

- Parâmetros de Ícones sonoros podem ser manipulados para interferir na percepção de urgência de uma mensagem sonora.

Como argumentado por Cao et al. (2010), pode-se concluir observando o conteúdo da revisão que é possível alterar os parâmetros do som de forma a torná-lo compatível com o nível de urgência de uma mensagem. Dentre esses parâmetros, estão os anteriormente comentados como: *pitch*, volume e ritmo que um som pode prover.

Em um teste, Fagerlönn, Lindberg e Sirkka, (2012) observaram que o intervalo entre os pulsos afetava a percepção dos participantes sobre a urgência da mensagem. O intervalo entre os pulsos de um ícone sonoro é o tempo de silêncio entre dois sinais sonoros, sejam eles o intervalo entre dois ícones sonoros de uma só nota, ou o intervalo entre duas notas de um mesmo ícone sonoro. Esse tipo de variação cria o efeito percebido como ritmo. Quanto menor o intervalo entre os pulsos, e consequentemente, mais rápido o ritmo, mais urgente era a percepção da mensagem.

Sobre esse mesmo parâmetro, Shahab, Terken e Eggen (2010) concluíram que se o intervalo entre os pulsos for variado, a duração dos pulsos deve ser a mesma, e que esse tipo de manipulação do pulso tem um efeito forte em como os participantes perceberam o nível de urgência das mensagens.

Além do ritmo, é possível explorar outros parâmetros sonoros para manipular a percepção de urgência. Politis, Brewster e Pollick (2014a) observaram que a variação da frequência é efetiva na criação de diferentes níveis de urgência para uma mensagem. Foi observado que quanto mais altas as frequências, maior era a percepção de urgência nas mensagens.

No entanto, é preciso cautela na criação de diferentes níveis de urgência, principalmente em situações nas quais a interpretação errônea de uma mensagem pode causar um acidente, como no trânsito. Foram observados resultados que apontam para uma tendência que diversos níveis de urgência apresentados por meio de ícones sonoros podem ser mal interpretados. Em um estudo de Cao et al. (2010), o experimento realizado comparou a eficácia do som e da vibração na apresentação de avisos de variados níveis de urgência. Resultados mostraram que os dois níveis extremos de urgência (menos e mais urgentes possível) foram melhor reconhecidos do que os níveis intermediários.

Ao observar esses resultados, pode-se concluir que os parâmetros das ondas sonoras podem ser manipulados para criar mensagens com variados níveis de urgência. Tal possibilidade é fundamental quando se trata de ADAS, uma vez que esse tipo de interface é totalmente focado em apresentar mensagens ao motorista sobre a tarefa principal de dirigir. Como diversas situações inesperadas podem acontecer em um contexto de condução, é crucial que a mensagem seja urgente e facilmente reconhecida, de forma a voltar toda a atenção do motorista para a situação de perigo.

- O aumento na noção de Urgência em ícones sonoros resulta em um aumento da percepção de irritabilidade

Aumentar parâmetros como *pitch* e ritmo pode ajudar a apresentar mensagens mais urgentes, porém, é preciso equilibrar o quanto esses parâmetros irão irritar o motorista. Notas mais altas com as quais o motorista não está acostumado podem gerar irritação por destoarem demais do ambiente sonoro, justamente por chamarem a atenção do motorista.

Politis, Brewster e Pollick (2014a) observaram que quanto mais urgentes eram as mensagens, mais irritados ficavam os participantes. Além disso, ao avaliar as preferências dos usuários, Wang, Li e Chen (2012) observaram que os ícones sonoros foram percebidos como mais irritantes em relação aos ícones auditivos, o que também pode ser interpretado como indicativo de que os parâmetros de ícone sonoro irritam o usuário.

É muito importante considerar o nível de irritabilidade gerado por um aviso sonoro. Isso é algo a se considerar porque mesmo que um sistema consiga passar eficientemente a urgência de suas mensagens, será prejudicial ao motorista se irritá-lo ou distraí-lo da tarefa principal, e pode até levar o condutor a desligar o sistema.

- Ícones auditivos têm uma melhor taxa de reconhecimento de uma mensagem em comparação aos Ícones sonoros

Diferente dos artificiais ícones sonoros, ícones auditivos usam sons gerados por ações do dia-a-dia para atribuir sentido a uma mensagem em interfaces. Por já estarem associados a gestos, ícones auditivos são mais facilmente reconhecíveis, ajudando o motorista a criar uma imagem da ação que gerou um determinado som.

Wang, Li e Chen (2012) montaram um experimento no qual os participantes deviam adivinhar a qual ação estavam ligados os ícones auditivos e os ícones sonoros apresentados. O resultado do teste mostrou que o reconhecimento das mensagens para ícones auditivos foi muito alto, ao contrário dos ícones sonoros, cujos significados foram frequentemente confundidos. Tais fatos estão de acordo com afirmações de Roginska (2013): “A interface (de ícones auditivos) tem apelo por se basear em sons que as pessoas escutam no dia a dia.”. (Roginska, 2013, p 345)

- A combinação dos ícones sonoros com uma ou mais modalidades de mensagem pode ser eficiente na apresentação de mensagens urgentes.

As modalidades visual, auditiva e tátil podem ser combinadas para apresentar mensagens de diferentes níveis de urgência ao motorista. Cao et

al. (2010) observaram que o reconhecimento das mensagens pela modalidade tátil foi mais acurado que os sons, o que é uma vantagem, uma vez que ícones sonoros podem ser ambíguos se mal projetados. Por outro lado, os sinais sonoros se mostraram mais rapidamente identificados do que as vibrações, além de conquistarem a preferência dos participantes.

Em Politis, Brewster e Pollick (2014a), um teste que avaliou as modalidades tátil, visual e auditiva, combinadas e isoladas, mostrou que quanto mais modalidades, mais urgentes e mais irritantes eram as mensagens, e, além disso, o uso das três modalidades conjuntas resultou em um melhor desempenho da tarefa primária. Os autores concluíram que o número de modalidades pode apresentar mais urgência sem sacrificar o entendimento da mensagem e com um nível de irritabilidade aceitável, uma vez que os níveis de percepção de irritabilidade não aumentaram na mesma proporção que a percepção de urgência.

Analisando os diferentes resultados sobre a multimodalidade sua eficiência é perceptível. Cada modalidade tem uma característica específica que pode ajudar a suprir as desvantagens e carências do som. Apesar de ser um campo que ainda deve ser estudado de forma mais aprofundada, a multimodalidade pode ser promissora, principalmente na apresentação de mensagens urgentes e importantes, como as providas de um ADAS, e que devem ser utilizadas com atenção.

5.2 Saída de Voz em ADAS

- O uso de níveis de urgência apresentados por voz pode ser efetivo, mas devem ser projetados com cautela

Pelas características dos comandos de voz, pouco se viu esse tipo de interação sonora ser aplicado a sistemas relacionados à tarefa principal de dirigir, como os ADAS. Por serem lentos, a utilização da saída de voz como meio de apresentação de mensagens urgentes e que requerem uma resposta rápida do motorista pode ser perigosa. Um aviso verbalizado pode demorar muito mais tempo do que um estímulo visual, como uma placa, ou até mesmo um sinal sonoro simples, como um ícone

sonoro ou um ícone auditivo. Entretanto, pode ser possível fazer uso dessa interação, principalmente em sistemas que integram ADAS e IVIS.

Sobre esse tópico, Politis, Brewster e Pollick (2014b) conduziram um experimento que usou avisos de ADAS e IVIS por saída de voz. Os avisos foram projetados com diferentes níveis de urgência, sempre colocando uma primeira palavra para esclarecer o quão urgente é a mensagem e o detalhe sobre a ação logo em seguida. Alguns parâmetros como palavras com variados níveis de urgência e um tom de voz com variação emocional foram usados para conseguir que a mensagem expressasse diferentes níveis de urgência. As respostas dos participantes comprovaram a efetividade do sistema proposto, e apontaram novamente para a tendência de que mensagens mais urgentes geram um aumento na percepção de irritabilidade.

Politis, Brewster e Pollick (2014b) compararam três diferentes modalidades de apresentação de informação ao motorista: Auditivo (Voz), Tátil, e Auditivo + Tátil. O uso da saída de voz combinada com vibrações em mesma frequência foi chamado de *Speech Tactons* (ou Ícones táteis de voz, em tradução nossa). Os resultados mostraram que a condição que os *Speech Tactons* obtiveram resultados de maior percepção de urgência, irritabilidade e eficiência de alerta. Resultados como esses estão alinhados com a discussão da categoria anterior, mostrando que o uso da multimodalidade foi efetivo na percepção de níveis de urgência e, conseqüentemente, na irritabilidade.

Levando-se em considerações as características dos comandos de voz, dos ícones sonoros e observando os resultados dos experimentos citados, pode-se concluir que, quando projetados usando tons de voz variados e determinadas palavras, o uso de saída de voz em avisos pode ser tão eficiente quanto os ícones sonoros na indicação de mensagens urgentes, com o benefício de um maior detalhamento da mensagem. No entanto, é preciso tomar cuidado com a irritabilidade que pode ser gerada pelo sistema, principalmente no uso de diferentes modalidades.

5.3 Saída de Voz em IVIS



- O uso da saída de voz em IVIS pode melhorar o desempenho da tarefa primária em relação a entradas manuais, porém, pode aumentar o tempo de realização de uma tarefa secundária.

A utilização da interação por voz foi dividida neste artigo em saída e entrada de voz, por serem campos amplos e com seus respectivos aspectos a serem explorados. No âmbito da saída de voz, uma das principais vantagens é, inegavelmente, o fato de essa modalidade não demandar recursos visuais, mantendo os olhos do motorista na via.

Ao conduzirem um experimento que testou a diferença entre o uso de um GPS tradicional (aparato visual combinado com saída de voz) e um modelo com *feedbacks* apenas por saída de voz, Kun et al. (2009) observaram na condição exclusivamente sonora o tempo de fixação visual na via foi maior. Ainda, mostrou que o GPS com aparato visual gerou efeitos negativos no desempenho da tarefa primária em relação à interface com *feedback* puramente sonoro.

Apesar do uso da saída de voz como meio de transmissão da mensagem ser benéfico à tarefa primária, a mesma tendência não é observada quando se trata da tarefa secundária. Weinberg, Harsham e Medenica (2011) conduziram um teste comparando o uso de três tipos de interfaces: *Head Down Display*, HDD, (Interfaces comuns, acopladas ao interior do carro), *Head Up Displays*, HUD, (interfaces projetadas no vidro frontal do automóvel) e a condição unicamente sonora, por saída de voz. O experimento revelou que, apesar da condição sonora ter mostrado melhor desempenho na direção e uma carga mental reduzida em relação às outras, o tempo de completude da tarefa secundária foi o maior dentre as condições. Entretanto, mesmo com a desvantagem de um aumento no tempo da tarefa secundária, os participantes mostraram em suas preferências subjetivas que se sentiram mais seguros com a condição auditiva, e que esse tipo de interação tem o mesmo nível de preferência que a HUD.

Esse tipo de comportamento pode apontar para uma tendência dos participantes de aceitarem as desvantagens dos sistemas auditivos para se manterem mais seguros. Como a informação

apresentada nos IVIS não é crucial para a direção, pode-se considerar aceitável um tempo maior para completar as tarefas, em troca de segurança. Apesar disso, é preciso minimizar esses efeitos ao máximo, como será discutido na próxima diretriz.

- O uso de um sistema que se adapta à situação da via pode ser benéfico para o desempenho da tarefa primária e pode atenuar as desvantagens do uso de saída de voz.

Como exposto anteriormente, um sistema adaptativo está ligado a capacidade de uma interface de se flexibilizar, seja às diferentes condições da via e do tráfego ou às necessidades do motorista. O benefício desse tipo de adaptação de um sistema é grande porque atenua desvantagens providas das interações sonoras, como o aumento do tempo de completude da tarefa. Alguns autores propuseram diferentes e promissoras abordagens nas interfaces sonoras para motoristas.

A proposta de Kennington et al., (2014) comparou duas interfaces de saída de voz e como elas afetam o desempenho das tarefas principal e secundária. Os autores propuseram uma interface adaptativa capaz de reconhecer os momentos nos quais o motorista precisa de mais atenção na via e capaz de interromper a saída de voz durante tal momento, com o objetivo de não distrair o condutor naquele período. Os resultados do teste mostraram um melhor desempenho da tarefa secundária com o sistema adaptativo e, não só foi observada uma melhora no desempenho da direção em relação a condição não-adaptativa, como também não foram observados impactos negativos na direção em relação à condição base (sem tarefa alguma). No entanto, a preferência dos usuários foi maior para a condição não-adaptativa, o que pode ser resultado de as interrupções parecerem com erros do sistema.

Similarmente, Macek et al., (2013) propuseram o uso de uma interface majoritariamente passiva. Artigo compara dois tipos de interfaces de reconhecimento de voz para tarefas secundárias: Passiva e Ativa. O sistema funcionava como um rádio. A informação é apresentada ao motorista de forma intercalada com músicas, e apresentava mensagens como níveis de gasolina ou notícias por meio de um ou mais “apresentadores” do sistema

do carro. Dessa forma, as necessidades por informação do motorista eram majoritariamente supridas de forma passiva, dando a ele, também, a opção de perguntar ao sistema sobre alguma informação eventualmente perdida. Foi observado que a interface passiva distraiu menos os motoristas da tarefa primária e obteve maior preferência dos participantes do que a ativa.

A capacidade de se adaptar às diferentes situações de um ambiente extremamente mutável como o trânsito pode ser difícil de ser alcançada por sistemas embarcados, mas com certeza é uma abordagem promissora no campo da interação sonora. Pode-se argumentar que quando a interface se adapta às condições do motorista, ela está reduzindo a necessidade de adaptação à interface, possibilitando a maior concentração na via, entendendo as necessidades dos usuários e abordando a forma menos intrusiva de atendê-los.

5.4 Entrada de Voz em IVIS

- Interfaces naturais de reconhecimento de voz trazem benefícios para as tarefas realizadas pelo motorista

As interfaces que se baseiam na interação por comando de voz requerem a entrada de um comando do motorista na maioria das vezes. No entanto, geralmente, os comandos necessários para a realização das ações do sistema são limitados a uma frase, palavras ou até mesmo uma ordem dos termos pré-estabelecidos. Porém, ao manipular a maneira com que tais comandos são reconhecidos pelo sistema, é possível chegar em uma interface natural de reconhecimento de voz, uma NLU (Natural Language Understanding). Esse tipo de aplicação pode trazer benefícios e mostra o quanto importante é o contexto na *design* interação.

Uma das vantagens desse tipo de interface é, como discutido por Winter; Grost e Tsimhoni (2010), o fato dos usuários não precisarem se lembrar dos diversos comandos necessários para realizar uma ação. Essa liberdade permite que os usuários realizem as ações sem se preocuparem em errar um termo específico do comando, gerando uma maior taxa de sucesso na realização das atividades. Segundo Labský et al. (2012), o aumento dos erros

no reconhecimento de voz pode gerar não só uma piora no desempenho da tarefa secundária, mas um aumento na carga mental provinda da frustração gerada pelas falhas do usuário.

No entanto, a eficiência e superioridade desse tipo de sistema em relação às interfaces comuns depende, como as outras interações, do contexto no qual ele está inserido. Características como a nacionalidade dos usuários também podem modificar os resultados da interface.

Hackenberg et al. (2013) conduziram um teste comparando, em três países diferentes, dois tipos de interfaces de reconhecimento de voz: a comum e a natural (NLU). O objetivo do estudo era verificar a efetividade e nível de eficiência das interfaces e as possíveis diferenças nos resultados na Alemanha, na China e nos Estados Unidos. Apesar do sistema NLU ter gerado menos distração e ter obtido melhores taxas de aceitação e reconhecimento, os níveis observados de melhora não foram iguais entre os três países.

As interfaces naturais e flexíveis de reconhecimento de voz conseguem tornar a interação entre o humano e o automóvel mais orgânica e instintiva, possibilitando, até um certo ponto, que o condutor use as palavras comuns ao seu vocabulário, preservando o sentido original do comando. Como observado nos testes, algumas considerações devem ser feitas para a criação desse tipo de interface, principalmente em relação ao entendimento das necessidades do usuário e das diferenças que os fatores culturais podem carregar.

5.5 Spearcons e Spindexes em IVIS

- O uso de Spindexes pode ser benéfico para ambas as tarefas primária e secundária na interação com listas e menus

Os *spindexes* geralmente representam a primeira letra de uma palavra e são mais comumente usados em listas e em menus, e são nesses tipos de interface que o benefício do seu uso aparece. A proposta de usar *spindexes* para acelerar o processo de navegação por voz em menus e listas se mostrou eficiente tanto para tarefas de secundárias, quanto para a própria condução.

Jeon et al. (2009), em um experimento, testaram a eficácia de diferentes tipos de menus auditivos, como *Text-To-Speech* (voz-para-texto) - TTS -, *Spearcons* e *Spindexes*, utilizados de forma isolada e combinada, em um contexto de multitarefas. Os resultados mostraram que o uso da condição de *Spindex* + TTS obteve uma redução no tempo de completude da tarefa secundária e a condição de *Spindex* reduziu a carga mental em relação às outras. Além disso, as abordagens de *Spindex* + TTS e *Spindex* + TTS + *Spearcon* alcançaram a maior taxa de preferência entre os participantes.

Apesar de recentes, *Spindexes* têm um grande potencial como interação por voz. *Spindexes* diminuem o tempo necessário para realizar uma tarefa secundária, mesmo quando o sistema não é flexível. As preferências dos usuários também apontam para uma maior satisfação quando são usados. Por serem conceitos novos, uma pesquisa mais aprofundada sobre o assunto é necessária.

6 Conclusão

O aumento no número de motoristas usando sistemas embarcados em seus carros vem aumentando, e, com eles, aumentam também os possíveis riscos nessa interação. A distração visual pode tirar os olhos do motorista da via, podendo causar pioras no desempenho da tarefa de condução e, eventualmente, acidentes no trânsito. Foi buscada, neste artigo, uma alternativa à interação pelo canal visual: a interação sonora.

Visando esclarecer as vantagens e desvantagens da interação sonora, foi feita uma revisão de literatura que identificou os desdobramentos do som como interação e gerou discussão sobre o tópico.

O som se mostrou um canal com muitas vantagens para a interação de motoristas com sistemas, principalmente por seus impactos na condução. Em sistemas de avisos como ADAS, parâmetros sonoros foram positivamente associados a percepção de urgência, levando a conclusão de que esse tipo de modalidade é eficiente em alertar o motorista sobre as diferentes e inesperadas situações na via. Além disso, a interação sonora em IVIS se mostrou eficiente na apresentação de informação sem a necessidade de um aparato

visual. Resultados mostraram que o uso de comando de voz nesses tipos de tarefas secundárias aumentou o tempo que o motorista passou olhando para a via. Ainda, medidas objetivas mostraram um melhor desempenho na tarefa de direção, que pode ser relacionada ao fato da atenção visual estar totalmente concentrada na condução.

Entretanto, alguns efeitos negativos foram observados na interação sonora. Primeiramente, em sistemas de avisos sobre a tarefa de direção, os alarmes com mais urgência causaram uma maior percepção de irritabilidade nos motoristas. Outro impacto negativo observado, dessa vez na tarefa secundária, foi um aumento no tempo de completude que alguns tipos de interação sonora causavam. A saída de voz, por sua característica demorada, prolongou tarefas paralelas à direção. Ainda no âmbito dos comandos de voz, a necessidade de termos pré-estabelecidos a serem lembrados e os possíveis erros decorrentes dessa interação podem levar à frustração do motorista e em um aumento da carga mental.

A relação entre a irritabilidade, carga mental e o desempenho de tarefas do motorista ainda precisa ser investigada, mas alguns resultados apontam para uma melhora no desempenho da direção quando a carga mental é mais baixa. Nesse sentido, além do notório benefício de não ocupar o canal visual, permitindo dedicação exclusiva dele à direção. Estudos como a Teoria de Múltiplos Recursos argumentam que a divisão dos recursos mentais em diferentes canais pode aliviar a carga mental gerada no motorista, apesar de não garantir que outros recursos de processamento ou de códigos não concorram entre si. Além disso, mesmo podendo ajudar a prever a carga mental que afetará o usuário em um contexto de múltiplas tarefas, a Teoria não evidencia qual das tarefas poderá ser afetada (Wickens, 2008).

Visando o abrandamento dos efeitos negativos gerados pelo som e a potencialização de suas vantagens, alguns campos promissores podem ser explorados. Eficiente na percepção de urgência, o uso de mais de uma modalidade - auditiva, tátil e visual - pode ser uma opção a ser acrescentada ao som. A baixa irritabilidade da modalidade visual e a maior taxa de reconhecimento da mensagem do

tato podem cobrir algumas das desvantagens dos alertas baseados em ícones sonoros, principalmente em condições nas quais as mensagens têm diversos níveis de urgência.

Outro campo promissor a ser explorado se relaciona diretamente à variação de prioridade nas informações apresentadas no veículo. Sistemas adaptativos que possam prever as necessidades do motorista para ajustar a maneira como a informação é exposta se mostraram eficientes no desempenho das tarefas primária e secundária, além de uma boa aceitação pelos participantes. A importância de saber quais tipos de interação usar e em quais momentos, balanceando resultados objetivos e as preferências dos usuários é evidente.

BIBLIOGRAFIA

AUGOYARD, J.-F. et al. **Sonic Experience**. Montreal: McGill-Queen's University Press, 2006.

CAO, Y. et al. **On timing and modality choice with local danger warnings for drivers**. In: Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '09. **Anais...**New York, New York, USA: ACM Press, 2009.

CAO, Y. et al. **Evaluating informative auditory and tactile cues for in-vehicle information systems**. In: Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '10. **Anais...**New York, New York, USA: ACM Press, 2010.

CARVALHO, L. R.; PEREIRA, A. T. C. User Experience Design and Sound Design : Dynamic Audio. **Ergodesign & HCI**, v. 2, n. 2, p. 1–10, 2014.

DASS, D. E.; UYTENDAELE, A.; TERKEN, J. **Haptic in-seat feedback for lane departure warning**. In: Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '13. **Anais...**New York, New York, USA: ACM Press, 2013.

FAGELÖNN, J.; LINDBERG, S.; SIRKKA, A. **Graded auditory warnings during in-vehicle use**. In: Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '12. **Anais...**New York, New York, USA: ACM Press, 2012.

RÖHLICH, P. et al. **Investigating safety services on the motorway**. In: Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '11. **Anais...**New York, New York, USA: ACM Press, 2011.

HAKENBERG, L. et al. **International evaluation of NLU benefits in the domain of in-vehicle speech dialog systems**. In: Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '13. **Anais...**New York, New York, USA: ACM Press, 2013.

HERMANN, T. et al. **Sonic Interaction Design**. In: HERMANN, T., HUNT, A., NEUHOFF, J. G. (Ed.). **The Sonification Handbook**. Berlin: Logos Publishing House, 2011. p. 87–110.

JEON, M. et al. **Enhanced auditory menu cues improve dual task performance and are preferred with in-vehicle technologies**. In: Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '09. **Anais...**New York, New York, USA: ACM Press, 2009.

KENNINGTON, C. et al. **Better Driving and Recall When In-car Information Presentation Uses Situationally-Aware Incremental Speech Output Generation**. In: Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '14. **Anais...**New York, New York, USA: ACM Press, 2014.

KUN, A. L. et al. **Glancing at personal navigation devices can affect driving**. In: Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '09. **Anais...**New York, New York, USA: ACM Press, 2009.

KUN, A. L.; SHYROKOV, A.; HEEMAN, P. A. **Spoken tasks for human-human experiments**. In: Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '10. **Anais...**New York, New York, USA: ACM Press, 2010.

LABSKÝ, M. et al. **Impact of word error rate on driving performance while dictating short texts**. In: Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '12. **Anais...**New York, New York, USA: ACM Press, 2012.

MACEK, T. et al. **Mostly passive information delivery in a car**. In: Proceedings of the 5th International

Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '13. **Anais...**New York, New York, USA: ACM Press, 2013.

MELOROSE, J.; PERROY, R.; CAREAS, S. **Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices**National Highway Traffic Safety Administration Docket No. NHTSA-2010-0053 Visual-Manual. [s.l: s.n.].

OMS. **Mobile phone use: a growing problem of driver distraction**. [s.l: s.n.]. Disponível em: <http://www.who.int/violence_injury_prevention/publications/road_traffic/distracted_driving/en/>.

ÖSTLUND, J. et al. **Driving performance assessment - methods and metrics**Contract. Linköping, Sweden: [s.n.].

POLITIS, I.; BREWSTER, S. A.; POLLICK, F. **Evaluating multimodal driver displays under varying situational urgency**. In: Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '14. **Anais...**New York, New York, USA: ACM Press, 2014.

POLITIS, I.; BREWSTER, S. A.; POLLICK, F. **Speech Tactons Improve Speech Warnings for Drivers**. In: Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '14. **Anais...**New York, New York, USA: ACM Press, 2014.

ROGINSKA, A. Auditory Icons, earcons and displays: Information and expression through sound. In: TAN, S.-L. et al. (Eds.). **The Psychology of Music in Multimedia**. [s.l.] Oxford University Press, 2013. p. 349–364.

SHAHAB, Q.; TERKEN, J.; EGGEN, B. **Auditory messages for speed advice in advanced driver assistance systems**. In: Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '10. **Anais...**New York, New York, USA: ACM Press, 2010.

TASHEV, I.; SELTZER, M. L.; JU, Y.-C. **Speech and sound for in-car infotainment systems**. In: Adjunct Proceedings of 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications. **Anais...** New York, New York, USA: ACM Press, 2009.

TERKEN, J.; VISSER, H.-J.; TOKMAKOFF, A.

Effects of speech-based vs handheld e-mailing and texting on driving performance and experience. In: Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '11. **Anais...**New York, New York, USA: ACM Press, 2011.

TUCHSCHEERER, S. et al. **Statistical effects of selected noise characteristics on speaker recognition in automotive environments**. In: Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '11. **Anais...**New York, New York, USA: ACM Press, 2011.

WANG, M. J.; LI, Y. C.; CHEN, F. **How can we design 3D auditory interfaces which enhance traffic safety for Chinese drivers?** In: Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '12. **Anais...**New York, New York, USA: ACM Press, 2012.

WEINBERG, G.; HARSHAM, B.; MEDENICA, Z. **Evaluating the usability of a head-up display for selection from choice lists in cars**. In: Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '11. **Anais...**New York, New York, USA: ACM Press, 2011.

WICKENS, C. D. Multiple Resources and Mental Workload. **Human Factors**, v. 50, n. 3, p. 449–455, 2008.

WINTER, U.; GROST, T. J.; TSIMHONI, O. **Language pattern analysis for automotive natural language speech applications**. In: Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '10. **Anais...**New York, New York, USA: ACM Press, 2010.

YANG, Y. et al. **Exploring differences in the impact of auditory and visual demands on driver behavior**. In: Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '12. **Anais...**New York, New York, USA: ACM Press, 2012.