

ALEXA: UMA ANÁLISE DE INTERFACE BASEADA EM VOZ

ALEXA: A VOICE-BASED INTERFACE ANALYSIS

Juliana Calhau PEREIRA

juliana.pereira99@outlook.com

Aluna do curso Bacharelado em Ciência da Computação, Centro Universitário Padre Anchieta – Jundiaí, SP

Carlos Eduardo CÂMARA

ccamara@anchieta.br

Professor do curso Bacharelado em Ciência da Computação, Centro Universitário Padre Anchieta – Jundiaí –
Doutor em Engenharia Elétrica – Teoria da Informação – FEEC-Unicamp

RESUMO

Tecnologias como Inteligência Artificial, Machine Learning, entre outras, possibilitam a criação das mais variadas técnicas e ferramentas que estão inseridas no nosso dia a dia. Dentre essas ferramentas, estão os assistentes pessoais inteligentes. Capazes de fazer consultas, criar rotinas, marcar calendários, entre muitas outras coisas, aprendem com o gosto do usuário utilizando comunicação por voz. Porém nem tudo é perfeito. Estes dispositivos podem não apresentar uma interação de voz compatível aos esperado e podem incomodar os usuários. Este artigo tem como objetivo verificar se a Assistente Virtual Alexa consegue realizar suas atividades requeridas ou ocorrem erros inesperados que possam vir a impedir a eficácia da interação entre homem-máquina, utilizando os conceitos das 10 Heurísticas de Nielsen.

Palavras-Chave

Alexa; Assistente Virtual; Heurística de Nielsen; Inteligência Artificial;

ABSTRACT

Technologies such as Artificial Intelligence, Machine Learning, among others, enable the creation of the most varied techniques and tools that are inserted in our daily lives. Among these tools are intelligent personal assistants. Able to make consultations, create routines, schedule calendars, among many other things, they learn with the user's discretion using their voice. However, not everything is perfect. These devices may not have a voice interaction compatible with those expected and may disturb users. This article aims to verify if the Alexa Virtual Assistant is able to perform its required activities or unexpected errors occur that may hinder the effectiveness of the interaction between man-machine, using the concepts of Nielsen's 10 heuristics.

Keywords

Alexa; Virtual Assistant; Nielsen's heuristic; Artificial intelligence.

INTRODUÇÃO

Uma vez, Arthur C. Clarke (16 de dezembro de 1917 – 19 de março de 2008), inventor, autodidata e autor britânico de livros de ficção científica, declarou em 1973: “Qualquer tecnologia suficientemente avançada é indistinguível de magia”. Também afirmou: “Os habitantes mais inteligentes do mundo no futuro não serão homens ou macacos. Serão máquinas. E, por fim, vão acabar superando seus criadores.” (GREELANE, 2017). Mesmo que ainda estejamos avançando pouco a pouco no desenvolvimento de máquinas, suas teorias não estão muito longe de se tornar realidade.

Desde carros autônomos, assistentes de voz, a Inteligência Artificial (IA) mudou a sociedade moderna, e foi assim que aconteceu com a voz. Por anos e de forma gradual, foi-se aperfeiçoando o conceito de computadores com os quais poderíamos conversar (COATES, 2019). Interfaces de voz do usuário (VUI) e até mesmo a Internet das Coisas (IoT) já existem há muitos anos, utilizando estes diversos conceitos. Estes sistemas são apoiados por aprendizado de máquina (ML), inteligência artificial (AI) e outros avanços, como a computação cognitiva, além do rápido avanço na minimização de componentes de hardware e o baixo custo dos mesmos com o avanço dos anos. Hoje em dia, desenvolvedores de software/hardware podem integrar voz em seus próprios dispositivos com muito pouco entendimento de processamento de linguagem natural (PNL), compreensão de linguagem natural (NLU), fala para texto (StT), texto para fala (TtS) e outras aplicações direcionadas ao mundo de interface de voz, de uma maneira simples e fácil (QUESADA, 2018). Mesmo com a popularidade do aplicativo Alexa (Figura[1]) e do Amazon Echo (Figura[2]), a Siri já estava em dispositivos iOS desde 2011, o Google Assistente(Figura[3]) está em telefones Android, de uma forma ou de outra, desde 2016, e a Cortana está incluída no Xbox e no Windows desde 2015 (COATES, 2019).

Figura 1- Aplicativo Alexa no smartphone

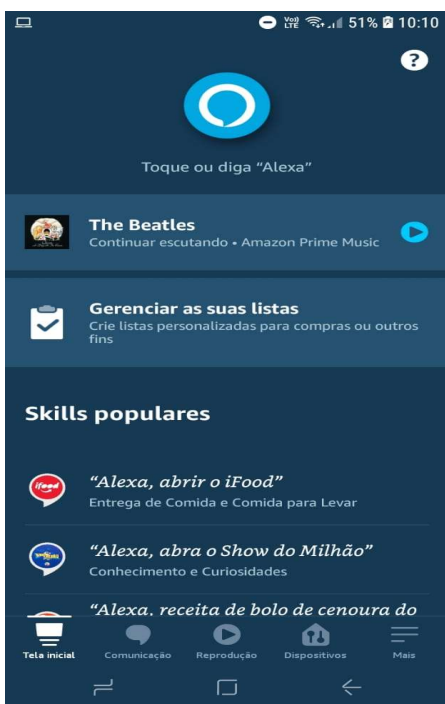


Figura 2- Amazon Echo



Figura 3- Aplicativo Google Assistente



Porém nem tudo aconteceu de uma forma tão rápida quanto vemos nos dias de hoje. Homer Dudley (engenheiro da Bell Labs) inventou a primeira máquina de síntese de fala, intitulada "Voder", durante o período da revolução técnico científica de 1937. Baseado em seu trabalho de 1928, o Vocoder (um codificador/decodificador de voz analógico), Dudley montou tubos de gás, um pedal, 10 teclas semelhantes às de piano e incluiu vários outros componentes projetados para trabalhar em harmonia na produção de fala humana a partir de som sintetizado. No ano de 1962, no Seattle World's (Justo), a IBM lançou a primeira máquina de reconhecimento de fala chamada "IBM Shoebox", que compreendia 16 palavras em inglês, o que foi impressionante para a época (QUESADA, 2018). As pesquisas continuaram nas décadas de 1960 e 1970, ampliando a quantidade de palavras que poderiam ser reconhecidas, permitindo o surgimento do reconhecimento de fala "contínua", não havendo necessidade de pausas entre as palavras. As pesquisas continuaram na década de 1980, na qual o reconhecimento de fala foi melhorado, e na década de 1990 surgiu o primeiro sistema independente de fala (qualquer pessoa poderia interagir com ele). A partir da década de 2000, os sistemas de resposta de voz interativa (IVR) fizeram grande sucesso, compreendendo a fala humana, bastando a pessoa ter apenas um telefone para solicitar cotação de valores, horários de filmes, transferência de dinheiro, informação de trânsito, entre outras coisas (PEARL, 2017).

Sistemas IVR permitiram uma evolução nos sistemas de fala, pois reconheciam *strings* longas (como o Fedex) e até mesmo falas mais complexas, como apostas em corridas de cavalos. Embora fossem uma inovação, os sistemas antigos eram mais simples do que os atuais, pois apenas acompanhavam o que já foi dito nos chamadores e acrescentavam informações, com base em predições do que seria dito pelo usuário (PEARL, 2017). Com o passar dos anos, as interfaces de voz se tornaram cada vez melhores, possibilitando que empresas desenvolvessem sistemas de compreensão de fala que funcionam tanto em dispositivos móveis como em dispositivos domésticos inteligentes que utilizam linguagem natural, como Google Assistant, Amazon Alexa, Siri e Microsoft Cortana (ALLIANNEJADI, 2020).

Em 2017, foi feita uma estimativa de que 35,6 milhões de pessoas iriam interagir pelo menos uma vez por mês com um dispositivo (SCIUTO, 2018). O que atrai o uso de dispositivos de voz e sua crescente adoção por parte de milhões de usuários é sua facilidade de uso. Usuários podem solicitar tocar uma música, pesquisar comandos na web, obter previsão do tempo, fazer compras online, criar tarefas e listas de compras, controlar dispositivos domésticos inteligentes utilizando apenas comandos por voz (SIEGERT, 2019). O usuário não precisa saber uma instrução ou decorar como deve ser a utilização, a simplicidade está junto à facilidade de

uso. Agora que milhares de pessoas podem comprar e configurar um agente de conversação inteligente por um baixo custo, é importante entender como as teorias dos pesquisadores e instituições combinam com a realidade de viver com dispositivos imperfeitos (SIEGERT, 2019).

Em seu artigo *The Paradox of Intelligent Assistants: Poor Usability, High Adoption*, 2018 (pág. 1), um cientista da computação com Ph.D. em Interação Homem-Máquina, Jacob Nielsen, descreve: “Os assistentes baseados em voz estão se tornando cada vez mais populares, 46% dos adultos dos EUA relataram usar controle por voz de assistentes digitais em 2017, de acordo com o Pew Research Center.”. Embora existam várias empresas desenvolvendo sistemas de voz, como Microsoft e Google, a Amazon controla 70% do mercado, com entre 7 e 11 milhões de pessoas que possuem pelo menos um dispositivo Alexa (SCIUTO, 2018). Essa é a razão de o enfoque deste trabalho ser essa interface de voz.

O objetivo deste trabalho é apresentar conceitos básicos de interfaces de voz e como as 10 Heurísticas propostas por Nielsen poderiam ser interpretadas para interfaces de voz; nota-se que apesar das expectativas relativas ao uso deste tipo de inteligência virtual, as interfaces de voz são ainda muito limitadas. Mesmo que este tipo de interação tenha evoluído muito ao longo do tempo, como dito anteriormente por Coates, a necessidade de testar e verificar a usabilidade nestas interfaces ainda é necessária. Por último, será proposto um teste onde será verificado se a interação verbal da Alexa está de acordo com as 10 Heurísticas de Nielsen, com base na visão da autora, e por fim serão descritas algumas possíveis soluções para problemas encontrados.

CONCEITOS BÁSICOS

Barbosa (2010) define alguns critérios básicos utilizados em interfaces: a interface com usuário determina os processos de interação possíveis, à medida que determina o que ele pode falar ou fazer, de que maneira e em que ordem. Portanto, quando definimos como a interação deve ocorrer, estamos restringindo ou determinando algumas características da interface, e vice-versa.

O contexto de uso influencia a interação de pessoas com sistemas interativos, pois elas estão inseridas em determinada cultura, sociedade e organização, possuem modo próprio de realizar suas atividades, possuem conhecimentos e concepções próprios e utilizam linguagem para interagir com as outras pessoas. É importante estarmos cientes de que o contexto de uso costuma ser diferente do contexto em que os desenvolvedores estão inseridos e com o qual estão acostumados. Daí a importância de investigarmos o contexto de uso com foco nos usuários e sob o seu ponto de vista. Isso nos permite avaliar o impacto dos diferentes aspectos do contexto sobre a interação humano-computador sendo concebida ou avaliada.

A usabilidade está relacionada com a facilidade de aprendizado e uso da interface, bem como a satisfação do usuário em decorrência desse uso (NIELSEN, 1993). Tradicionalmente, a usabilidade enfoca a maneira como o uso de um sistema interativo no ambiente de trabalho é afetado por características do usuário (sua cognição, sua capacidade de agir sobre a interface e sua capacidade de perceber as respostas do sistema). Com a disseminação dos sistemas computacionais interativos em ambientes diferentes do trabalho, a usabilidade passou a englobar também as emoções e os sentimentos dos usuários.

Para um usuário tirar proveito do apoio computacional oferecido pelo sistema, não podem existir barreiras que o impeçam de interagir com sua interface. O critério de acessibilidade está relacionado à remoção das barreiras que impedem mais usuários de serem capazes de acessar a interface do sistema e interagirem com ele. Cuidar da acessibilidade significa permitir que mais pessoas possam interagir com o sistema, tenham elas alguma deficiência ou não. A intenção é incluir, não excluir. O critério de acessibilidade está relacionado com a capacidade de o usuário acessar o sistema para interagir com ele, sem que a interface imponha obstáculos.

O critério de comunicabilidade chama atenção para a responsabilidade de o designer comunicar ao usuário suas intenções de design e a lógica que rege o comportamento da interface. Esse critério se pauta no pressuposto de que se o usuário tiver acesso à lógica de design, ele terá melhor condição de fazer um uso produtivo e criativo do apoio computacional oferecido pelo sistema.

ESTADO DA ARTE

Segundo Silva (2019), agentes conversacionais de voz são classificados como dispositivos tecnológicos que interagem com o ser humano por meio da inteligência artificial, permitindo que este dispositivo seja capaz de detectar som e reagir a estímulos externos.

A Interface Conversacional permite uma interação bilateral, onde por vezes o homem fornece informações ao computador, outras vezes é o computador que fornece informações ao homem. Quando se pensa neste tipo de interface, há um cenário mais complexo, pois este tipo de modelo pode sofrer alterações mais significativas. O usuário precisa sentir que suas ações e passos na navegação são naturais.

Um dos aspectos que desafiam os estudos e o desenvolvimento de interfaces é em relação a como apresentar ao usuário estímulos e uso a partir de um recurso simples, como manter o ícone do aplicativo presente quando comparado a uma interface visual. Estes desafios se evidenciam cada vez mais em pesquisas, na tentativa de desenvolver uma experiência que não seja irritante para o usuário nem que torne o assistente alguém que fale demais.

Cruz (2013) discorre que estes assistentes inteligentes são o resultado do avanço da pesquisa na área de ciência da computação que combina diversas áreas, dentre elas:

Inteligência Artificial: permite realizar atividades não apenas repetitivas, numerosas e manuais, como também as que demandam análise e tomada de decisão. Desta forma, exibe características similares às dos seres humanos, permitindo o aprendizado e o avanço cognitivo.

Processamento de Linguagem Natural: subárea da Inteligência Artificial (IA) que estuda a capacidade e as limitações de uma máquina em entender a linguagem dos seres humanos. O objetivo do PLN é fornecer aos computadores a capacidade de entender e compor textos. Permite que os usuários usem linguagem natural para se comunicar com os assistentes.

Banco de Dados: permite que a máquina armazene e recupere grandes volumes de informações. Hoje, existem diversos tipos de SGBDs, e cada um é adequado para uma necessidade dos clientes.

Redes de Comunicação de Dados: permite que os Assistentes Virtuais estejam nas páginas a serem acessadas, com computadores pessoais, tablets e smartphones.

Descreve também que os assistentes virtuais podem ter diversas aplicações. Dentre elas estão:

Saúde e bem-estar: com base nos dados dos usuários, os assistentes podem indicar dietas, exercícios, hábitos de consumo, além de informações personalizadas. Também podem indicar profissionais da saúde que estejam em conformidade com as necessidades dos usuários. Podem também ofertar a compra de equipamentos para prática de exercícios físicos, além de suplementos alimentares.

Comércio Eletrônico: além de auxiliarem na compra e venda de mercadorias pela internet, solucionam dúvidas sobre características, preços e prazos de entrega, com base nos requisitos dos clientes. Permitem mostrar ofertas personalizadas de acordo com o perfil do cliente e com o momento. Também podem ser usadas para solicitar reservas de hotel, carro, vagas, restaurantes, entre outros, gerenciando cancelamento e agendamento destas reservas, receber reclamações ou sugestões e realizar recuperação de senhas e logins necessários na maior parte dos sites.

Educação: permite que os estudantes, por meio do E-learning, possam tirar dúvidas e ser direcionados nos estudos por Assistentes Virtuais. Esses Assistentes adquirem conhecimento sobre as melhores opções de direcionamento e respostas conforme sua utilização. Além de serem utilizados para sanar dúvidas sobre matemática, física, química, geografia, história etc., possibilitam que os alunos tenham sempre um auxílio inesgotável.

Turismo: ajudam a encontrar os melhores pacotes turísticos e a esclarecer dúvidas sobre documentos e requisitos para cada tipo de roteiro, com base nas solicitações dos usuários. Podem também informar os horários de chegada e saída dos mais variados tipos de transporte, as condições de tráfego e o clima. Esses assistentes têm um objetivo em comum: ser eficiente 24h por dia.

Medicina: capazes de assimilar grandes bancos de dados em sua memória, os Assistentes Virtuais podem tirar dúvidas de médicos e pacientes sobre questões relacionadas a indicações e restrições para exames e medicamentos, marcação de consultas, leitos, resultados de exames, gestão de pacientes com doenças graves ou crônicas, auxiliando na cura e tratamento das mais diversas enfermidades.

Auxílio Doméstico: controlar determinados equipamentos inteligentes compatíveis com interação por voz.

Finanças: ajudam a esclarecer dúvidas acerca de opções de investimento e procuram selecionar a melhor opção possível para usuários. Também podem auxiliar na economia de suas reservas e de suas finanças. Podem determinar o perfil do cliente e selecionar a melhor opção disponível de investimento dentre as disponíveis.

Tipos de assistentes

Cruz (2013) observa que os assistentes virtuais inteligentes podem ser classificados por tipos, que variam conforme as características que cada um possui. São classificados conforme:

Comportamento

Passivos: apresentam-se apenas quando o usuário solicita alguma informação.

Dinâmicos: apresentam-se ao usuário assim que o mesmo se torna ativo dentro de um sistema, como site, smartphone, tablet etc.

Dinâmicos com Gatilho: apresentam-se quando o cliente aparenta, pelo seu comportamento, necessitar de ajuda. Por exemplo, após múltiplas tentativas falhas de fornecer uma senha para uma área restrita específica de um site ou quando o cliente retorna muitas vezes a mesma página sem motivo aparente.

Propósito

Gerais: auxiliam o usuário com assuntos de caráter geral, como a troca de informações sobre notícias que estão no meio de comunicação.

Especializados: auxiliam o usuário com assuntos específicos, tais como a compra de um veículo, informações financeiras, compras na internet, turismo etc.

Apresentação

Com Avatar: apresentam-se personificados na forma de uma imagem, normalmente na forma de uma figura humana ou um robô, podendo até mesmo ser estilizada.

Sem Avatar: não apresentam personificação nem imagens. Apresentam-se normalmente como uma caixa de diálogo com mensagens do tipo: "Como posso ajudar?".

Comunicação

Sociáveis: o Assistente se mostra atencioso e cortês, demonstrando certa preocupação com as informações que oferece. Tem uma direção mais personificada.

Indiferentes: o Assistente responde às questões de forma mecânica e indiferente ao nível da informação.

Integração

Integrados: são capazes de acessar os sistemas de informação corporativos para fornecer informações ao cliente. Também podem usar os dados fornecidos pelo cliente para atualizar as informações contidas nesse sistema.

Não Integrados: não são capazes de acessar os sistemas de informação e por isso são mais limitados nas informações que podem fornecer ao cliente.

Por que os assistentes têm voz feminina?

Uma das possíveis respostas da Siri, assistente pessoal inteligente da Apple, quando perguntamos qual é o seu sexo é: "A minha voz pode soar como uma mulher, mas eu existo além do conceito humano de sexo". Mesmo que ela se identifique como um organismo genderless (que não se identifica com nenhum sexo), a voz que chega aos nossos ouvidos quando ela responde a alguma solicitação é distinguida como feminina (FIORETTI, 2018). Porém não é só a Siri que possui esta voz. Alexa da Amazon, Cortana da Microsoft, e o Google Assistant *também tem claramente uma voz de mulher. E não é só isso. O nome também é de mulher. O nome da Siri, assistente da Apple, significa "linda mulher que te leva à vitória" na língua nórdica. Além disso, a escolha também foi influenciada por ser fácil de soletrar e de ser dita pelos consumidores* (FIORETTI, 2018).

No caso da Alexa, por mais que Amazon afirme que é a abreviação de Alexandria (biblioteca incendiada pelas tropas de Júlio César em 48 a.C), também é um nome feminino. No caso da Google, embora sua assistente digital não tenha um nome feminino, a voz que nos responde é a de uma mulher, e a cada atualização fica cada vez mais similar à voz humana (Fioretti, 2018). Mas não é à toa que estes assistentes costumam ter vozes femininas. De acordo com Karl MacDorman, professor da Universidade de Indiana e especialista na interação humano-computador, as pesquisas indicam que há uma maior aceitação do discurso feminino.

O professor Clifford Nass, autor de *Wired For Speech*, numa entrevista à CNN também afirmou: "*É muito mais fácil encontrar uma voz feminina que agrada a todos do que uma voz masculina, afinal, o cérebro humano é programado para gostar de mulheres*". Nass ainda completou explicando que, além de impulsionado pelo condicionamento social, *queremos que nossas assistentes virtuais sejam amigáveis e não ameaçadoras; prestativas, mas não dominantes. E isso acontece com voz feminina. Também para a maioria dos seres humanos, a voz feminina está diretamente relacionada ao papel materno, que representa no nosso subconsciente acolhimento, cuidado e segurança – o que de certa forma perpetua tal preferência. Porém, essa reflexão não é totalmente conclusiva. Por outro lado, um estudo publicado na revista Live Science dizia que para um sistema telefônico automatizado, a voz masculina era mais "utilizável", mas não necessariamente tão "confiável" quanto uma voz feminina* (Fioretti, 2018).

Podemos concluir que humanizar as assistentes pessoais com nomes, emoções, personalidades e gêneros (principalmente feminino) ajuda a construir a relação de confiança com os usuários, para que se sintam à vontade para interagir com este tipo de inteligência artificial (Fioretti, 2018).

As 10 Heurísticas de Nielsen e a proposta de Mittal

Pensar em UI (user interface) design é pensar em projetar uma interface que não gere momentos de insegurança para o usuário, que deixe claro quais serão os resultados de suas ações e garantir que o mesmo realize todas as tarefas de forma simples e eficiente; em outras palavras, fazer com que o usuário “não necessite de um manual de instruções”. Para isso existem dez heurísticas criadas por Jakob Nielsen, que ajudam a projetar uma boa interface e por consequência uma ótima experiência de uso (MACEDO, 2017). Mittal (2020) explica como funciona cada heurística proposta por Nielsen e logo após ela diz como isso poderia ser aplicado em uma VUI.

1- Visibilidade do status do sistema

O sistema deve sempre manter os usuários informados sobre o que está acontecendo, por meio de feedback apropriado dentro de um tempo razoável. No caso de uma VUI, o usuário deve saber quando o sistema está ouvindo, quando está processando e quando está falando. O sistema também deve informar os usuários quando algo está errado. Dicas visuais (como um anel de luz animado no Alexa Echo) e áudio não verbal (como fones de ouvido) podem ajudar na comunicação do status do sistema.

2- Combinação entre o sistema e o real mundo

O sistema deve falar a língua dos usuários, com palavras, frases e conceitos familiares para o usuário, em vez de sistemas orientados a termos. No caso de VUI, os humanos falam intuitivamente, mesmo sem perceber a complexidade de sua linguagem, mas as máquinas precisam levar em conta esta complexidade. Os sistemas de voz devem ser treinados para entender as regras básicas de conversação.

3- Controle e liberdade do usuário

Os usuários frequentemente escolhem funções do sistema por engano e precisarão de uma opção claramente marcada como "saída de emergência" para deixar o estado indesejado, sem ter que passar por um diálogo extenso. Apoio, suporte, desfazer e refazer. No caso da VUI, os sistemas de voz devem permitir que os usuários saibam quando eles dizem algo por engano e querem sair ou iniciar a correção.

4- Consistência e padrões

Os usuários não devem se perguntar se palavras diferentes, situações ou ações significam a mesma coisa. Numa VUI, embora ter um vocabulário consistente ajude as GUIs (Interface Gráfica do Usuário, em português), os usuários vão começar a ficar frustrados se as VUIs repetirem a mesma frase toda vez. Esta é a razão pela qual o sistema de voz deve entender a entrada que é expressa de muitas maneiras alternativas.

5- Prevenção de erros

Ainda melhor do que boas mensagens de erro, é um design cuidadoso que evita que um problema ocorra no primeiro local. Ou elimine condições propensas a erros, ou verifique-as e apresente aos usuários com uma opção de confirmação antes de se comprometerem com a ação. No caso de VUI, para evitar erros, o

sistema de voz deve confirmar sua hipótese com os usuários antes de tomar decisões importantes (como pedir algo sem confirmar o pedido).

6- Reconhecimento em vez de recordação

Minimize a carga de memória do usuário criando objetos, ações e opções visíveis. O usuário não deve ter que se lembrar de informações de uma parte do diálogo para outra. Instruções para uso do sistema devem ser visíveis ou facilmente recuperáveis sempre que apropriado. No caso de VUI, por causa da natureza efêmera da fala, este valor é 3 + 1. Para ajudar a focar a atenção do usuário no que é importante, novas informações devem ser colocadas no final ou próximo ao final da frase.

7- Flexibilidade e eficiência de uso

Aceleradores – invisíveis para o usuário novato – muitas vezes podem acelerar a interação para o usuário especialista, de modo que o sistema possa atender bem usuários inexperientes e experientes. Permitir que os usuários personalizem ações frequentes.

VUIs podem ter atalhos também para pular uma mensagem de boas-vindas a alguém que usa o sistema mais de cinco vezes ao dia e evitando confirmações. A recuperação de erros pode ser tratada de maneira diferente para iniciantes e especialistas. Para usuários avançados, novas solicitações rápidas (não fornecendo instruções detalhadas sobre o que o usuário deve dizer imediatamente) podem funcionar melhor do que escalar detalhes (fornecendo exemplos do que o usuário deve dizer).

8- Design estético e minimalista

Os diálogos não devem conter informações irrelevantes ou raramente necessárias. Cada unidade extra de informação em um diálogo compete com as unidades relevantes de informação e diminui sua visibilidade relativa. Para VUI, Paul Grice definiu quatro regras básicas de conversação cooperativa:

Qualidade: não diga aquilo para o que faltam provas;

Quantidade: não seja mais ou menos informativo do que o necessário;

Relevância: diga apenas coisas relevantes para o tópico;

Maneira: seja breve, vá direto ao ponto e evite ambiguidades e obscuridade.

O design estético e minimalista nas VUIs reside na clareza, bem como na brevidade dos diálogos sendo entregues.

9- Ajude os usuários a reconhecer, diagnosticar e recuperar de erros

As mensagens de erro devem ser expressas em linguagem simples (sem códigos), indicar precisamente o problema e sugerir construtivamente uma solução. Numa VUI, para lidar com erros de reconhecimento, os sistemas de voz devem ter estratégias robustas de recuperação de erros para lidar com eles. Eles devem ser claros o máximo que puderem ao informarem aos usuários o que deu errado (isso remonta à heurística 1, visibilidade do status do sistema) e o que eles podem fazer para continuar avançando na conversa.

10- Ajuda e documentação

Mesmo que seja melhor se o sistema puder ser usado sem documentação, pode ser necessário fornecer ajuda e documentação. Essas informações devem ser fáceis de pesquisar, com foco na tarefa do usuário, listar as etapas concretas a serem realizadas e não ser muito grande. No caso de VUI, se houver mais de três erros

em uma linha, as estratégias de recuperação não estão funcionando para trazer o usuário de volta à pista, o sistema deve dar a eles "você pode dizer/perguntar" mensagens. Ele pode fornecer ajuda escalonada, informando os usuários sobre a documentação a que eles podem se referir e o exemplo de interação do aplicativo.

TESTE/ANÁLISE

Nesta seção será documentada uma proposta de teste real em que, a partir dos conceitos apresentados por Mittal (2020), será aplicada a interface Alexa. O objetivo principal é verificar como a interface de voz Alexa (da Amazon) se comporta perante determinadas perguntas feitas a ela e se o resultado é compatível com as propostas desta autora.

A escolha da Alexa se deve a sua popularidade, maior quantidade de usuários e um estudo mais amplo da autora nesta interface. A escolha como aplicativo de celular também é determinada pela facilidade de as pessoas baixarem este aplicativo gratuitamente, diferente do Echo Dot ou Echo Show, que utilizam também a Alexa, porém é um dispositivo à parte, geralmente utilizado para interação e comunicação de dispositivos inteligentes.

Um teste de interface de voz poderia ser aplicado de duas formas: quando um desenvolvedor de software testa a aplicação, sendo necessário conhecer a linguagem de programação utilizada e a PLN (Processamento de Linguagem Natural), que consiste no desenvolvimento de modelos computacionais para a realização de tarefas que dependem de informações expressas em alguma linguagem natural. Desta forma, a PLN estuda construções de fonemas, sons e significado que permitem criar as interações numa interface de voz. O desenvolvedor saberia como é usada a construção da PLN e seria proposto um teste com base nestes conceitos e nas supostas questões que esta interface poderia resolver. (PEREIRA,1997)

A segunda opção de teste seria com um ou vários usuários. Este tipo de teste geralmente oferece perguntas, ações a serem feitas, e o usuário fala com a interface e dá o feedback do seu ponto de vista referente ao que foi proposto. Não é exigido conhecimento prévio da interface, programação ou PLN, mas apenas utilizar o seu conhecimento e aprendizado de usuário final e dar seu ponto de vista, positivo ou negativo, se teve dificuldades, facilidades, apontar melhorias etc. Este é o tipo de teste proposto neste documento, porém apenas sob o ponto de vista da autora como uma usuária comum.

Metodologia

Para o início da proposta, foi decidido que os testes serão apenas no dispositivo smartphone, utilizando o aplicativo Alexa, e somente a autora irá testar as questões. A escolha das questões se baseou em estudos dos materiais apresentados e outras fontes de consulta, além do entendimento pessoal da autora sobre o que um usuário comum poderia perguntar.

Para a interpretação das heurísticas de Nielsen, que são focadas especialmente para interfaces visuais, foi um pouco difícil pensar em como isso seria traduzido para interfaces de voz, nas quais o usuário não consegue perceber ou visualizar determinadas funções que apenas uma interface visual conseguiria transmitir.

Uma boa analogia seria a experiência com um site. A pessoa pode entrar em um determinado site, visualizar as informações que deseja, pesquisar, receber feedback interativo pelo qual consegue enxergar as respostas de erro, aviso, direcionamento, entre outros. Caso o usuário cometa alguma ação errada, há várias maneiras de fazer com que ele compreenda o ocorrido e seja guiado de forma correta.

Para uma interface de voz, isto é ainda mais complicado. O fato de um usuário utilizar apenas voz, sem contato visual, pode induzi-lo a determinados erros que ele não sabe distinguir se foram seus, da interface ou do dispositivo, e mesmo que descubra o que ocorreu, pode não saber como resolver.

Foi criada uma ficha do produto a testar que sintetizasse a primeira fase do processo de teste, que é a escolha do aplicativo a ser testado, sistema operacional e dispositivo.

Quadro 1. Descrição do produto do Ambiente de Uso, de Estudo e Teste do App Alexa (Amazon)

| NOME | SMARTPHONE |
|-------------------------------|---------------------------|
| Marca | Samsung |
| Modelo | J4 Plus |
| Sistema Operacional | Android Oreo versão 8.1.0 |
| Aplicativo de teste instalado | Alexa Amazon |
| Versão do Aplicativo de teste | 2.2.372932.0 |

Preparação e questões

A preparação dos testes envolveu a escolha de perguntas consideradas simples e outras questões mais complexas, pois o objetivo é saber como ela as interpreta. Envolveu o estudo e a análise de artigos sobre assistentes virtuais digitais, heurística de Nielsen, testes de usuários com dispositivos inteligentes, internet das coisas e computação. A preparação consistiu em:

Selecionar os materiais relevantes já apresentados e leitura atenta dos mesmos;

Planejar a escrita das possíveis perguntas;

Planejar possíveis problemas que poderiam ser encontrados em uma interface de voz;

Determinar em qual dispositivo seriam feitos os testes da aplicação;

Escrever as perguntas selecionadas com base em conversas já feitas anteriormente com a interface de voz.

Desta forma, foram escolhidas 10 questões, sendo avaliado se determinada pergunta infringe alguma regra das heurísticas. As questões propostas foram:

Q1: Alexa, qual a temperatura hoje em Jundiáí?

Q2: Alexa ligue para pai.

Q3: Alexa, informe as notícias do dia.

Q4: Alexa, preciso que o alarme toque daqui a 15 minutos.

Q5: Alexa, pesquise imagens de cachorros./ Alexa, pesquise no Google imagens de cachorros.

Q6: Alexa, pesquise sobre cachorros de porte médio.

Q7: Alexa, me diga quais as melhores faculdades de Minas Gerais, exceto as particulares.

Q8: Alexa, lembre-me de que tenho médico às 14 horas. /Alexa, cancele o médico às 14h.

Q9: Alexa, toque músicas do Spotify.

Q10: Alexa, há algum hotel no bairro Vila Nambi em Jundiáí?

RESULTADO DOS TESTES

Os testes foram feitos no dispositivo apresentado anteriormente, utilizando a assistente virtual Alexa. A instalação do aplicativo foi fácil, demorou menos de 10 minutos, e o login no aplicativo foi rápido, bastando já ter uma conta cadastrada na Amazon.

Foi necessário também um conhecimento do aplicativo; algumas conversas com o aplicativo foram feitas antes de iniciar os testes. Aprender as configurações do aplicativo possibilita entender que dependendo das respostas apresentadas, isto influenciaria na satisfação sobre a interface.

Para executar os testes e registrar os resultados, foram criadas tabelas no Word. A primeira, tabela1, mostra o script da frase testada, a resposta da interface e uma observação. Nas tabelas 2 e 3, há a análise de como a resposta da questão interage junto a cada uma das heurísticas. As tabelas são constituídas das seguintes colunas:

Tabela 1:

Resultado das questões feitas à Alexa

ID: identificação do script testado

Script: questão testada na interface de voz

Resposta da Interface de voz: resposta dada pela interface de voz Alexa

Observações: informações referentes à resposta e ao que era esperado

Tabelas 2 e 3:

Heurística: nome da heurística

Q: refere-se à questão analisada

Data dos testes: 16/11/2020 a 20/11/2020

Quantidade de tentativas por script: 3

Depois de definidas estas tabelas no Word, procedeu-se à realização dos testes. As respostas foram anotadas nas tabelas e registradas conforme descrito no documento abaixo.

Tabela 1. Resultado das questões feitas à Alexa

| ID Questão | Script | Resposta da Interface de voz | Observações |
|-------------------|--|--|--|
| 1 | Alexa, qual a temperatura hoje em Jundiaí? | Temperatura atual em Centro SP é de 24 graus Celsius com máxima de 26 graus. | Compreendeu a pergunta, porém exibiu a temperatura de outra cidade(?). Exibiu resposta também em tela. |
| 2 | Alexa, ligue para pai | Pai não está cadastrado para chamadas Alexa. Para ver para quem eu posso ligar, toque no ícone de bate-papo na parte de inferior e depois no ícone de lápis na parte superior. | Compreendeu a pergunta e orientou como cadastrar um contato no aplicativo Alexa. Usou apenas voz. |

| | | | |
|---|--|--|--|
| 3 | <p>Alexa, informe as notícias do dia.</p> <p>Próxima notícia.</p> <p>Alexa, pare.</p> | <p>Juliana, aqui estão as novidades:...</p> <p>(leitura das notícias)</p> <p>Aqui está a última notícia.</p> <p>Parou de falar</p> | <p>Informou as notícias disponíveis.</p> <p>Leu direto a última notícia.</p> |
| 4 | <p>Alexa, preciso que o alarme toque daqui a 15 minutos.</p> <p>Alexa, ligue alarme em 20 minutos.</p> <p>Alexa, cancele o alarme.</p> | <p>15 minutos começando agora.</p> <p>20 minutos começando agora.</p> <p>Seu alarme não foi configurado.</p> | <p>Configurou alarme.</p> <p>Não reconheceu/compreendeu o cancelamento do alarme.</p> |
| 5 | <p>Alexa, pesquise imagens de cachorros.</p> <p>Alexa, pesquise no Google imagens de cachorros.</p> | <p>Este dispositivo não é compatível com fotos pessoais.</p> <p>Sorry, I have a trouble.</p> | <p>Não possui compatibilidade com foto.</p> <p>Retornou uma resposta em inglês. Não compreendeu/não sabe a resposta.</p> |
| 6 | <p>Alexa, pesquise sobre cachorros de porte médio.</p> | <p>Sorry, I have a trouble.</p> | <p>Não compreendeu/não sabe a pergunta e retornou a resposta em inglês.</p> |
| 7 | <p>Alexa, me diga quais as melhores faculdades de Minas Gerais, exceto as particulares.</p> | <p>Ainda não posso olhar informações dos estabelecimentos.</p> | <p>Não sabe/Não informa o motivo pelo qual não pode consultar as informações.</p> <p>(Deve ser a pandemia?)</p> |

| | | | |
|----|--|---|--|
| 8 | Alexa, lembre-me de que tenho médico às 14 horas. Alexa, cancele o medico às 14h. | Certo, lembrarei você às 14h. Hum, estou com dificuldade. | Informou que configurou o lembrete. Não compreendeu/não sabe sobre a solicitação anterior. Talvez não saiba cancelar o médico. |
| 9 | Alexa, toque músicas do Spotify. | Para usar o Spotify, é necessário fazer download primeiro. Redirecionando você para a loja de aplicativos agora. | Informa que não possui o Spotify instalado e direciona à Playstore. |
| 10 | Alexa, há algum hotel no bairro Vila Nambi em Jundiaí? Alexa, pule para o próximo item. Alexa, pare. | Aqui estão algumas informações que consegui encontrar na Vila Nambi em Jundiaí. Exibiu uma lista com seis opções de hotéis/pousadas. Hum, algo deu errado. Parou de ler. | Exibiu uma lista de hotéis na tela e leu as opções. Parou a leitura, porém não consegue voltar a ler ou avançar as opções. |

*No Script 1, a autora se encontrava na cidade de Campo Limpo Paulista (SP) e tinha permitido a localização da região onde se encontrava durante o uso. Para todas as questões foi usada esta localização.

Estas foram as questões e as respostas envolvidas na proposta. Abaixo serão mostradas duas tabelas. Por questão de possuir muitas informações, a tabela teve de ser dividida em duas. Elas mostram as 10 Heurísticas de Nielsen e como cada questão e resposta dada da tabela anterior se conecta à heurística e à análise sobre ela.

Tabela 2. Análise correspondente às Heurísticas

| RESULTADOS | | | | | |
|----------------------------------|--|-------------------|-------------------|-------------------|--|
| Heurística | Q1 | Q2 | Q3 | Q4 | Q5 |
| Visibilidade e Status do Sistema | Status do sistema exibido corretamente | Status do sistema | Status do sistema | Status do sistema | Status do sistema exibido corretamente |

| | | | | | |
|--|--|---|--|---|--|
| | | exibido corretamente | exibido corretamente | exibido corretamente | |
| Combinação entre o sistema e o mundo real | Não deu informação pertinente à região | Compreendeu corretamente | Compreendeu corretamente | Não compreendeu a solicitação de cancelamento | Não compreendeu/Não sabe pesquisar na internet |
| Liberdade e controle do usuário | Script não exige ação para esta heurística | Script não exige ação para esta heurística | Permitiu corretamente a ação requerida | Não permitiu que fosse cancelado o alarme | Script não exige ação para esta heurística |
| Consistência e padrões | Script não exige ação para esta heurística | Script não exige ação para esta heurística | Script não exige ação para esta heurística | Compreendeu corretamente | Não permitiu/não entendeu entrada semelhantes |
| Prevenção de erro | Script não exige ação para esta heurística | Script não exige ação para esta heurística | Script não exige ação para esta heurística | Permite uma confirmação implícita, porém não preveniu como cancelar | Não se aplica |
| Reconhecimento em vez de recordação | Resposta simples e fácil de lembrar | Orienta a configuração no aplicativo, porém difícil de lembrar | Script não exige ação para esta heurística | Resposta simples e fácil de lembrar | Script não exige ação para esta heurística |
| Flexibilidade e eficiência de uso | Resposta simples e satisfatória | Usuários mais leigos terão dificuldade em fazer o que ela orienta | Fácil de ouvir as notícias | Fácil de configurar, porém difícil de cancelar | Há dificuldade em saber o que deu errado |
| Design estético e minimalista | Resposta simples e satisfatória | Resposta simples e satisfatória | Resposta simples e satisfatória | Resposta simples | Resposta confusa |
| Ajude os usuários a reconhecer, diagnosticar e recuperar erros | Não sabe que ocorreu um erro | Orientou como solucionar o problema | Script não exige ação para esta heurística | Não orientou como cancelar o alarme | Não orientou nem solucionou o problema |
| Ajuda e documentação | Script não exige ação para esta heurística | Não possui outro auxílio além da voz | Script não exige ação para esta heurística | Não fornece ajuda nem documentação | Não fornece ajuda nem documentação |

Tabela 3. Análise correspondente às Heurísticas

| RESULTADOS | | | | | |
|--|--|--|---|--|--|
| Heurística | Q6 | Q7 | Q8 | Q9 | Q10 |
| Visibilidade e Status do Sistema | Status do sistema exibido corretamente | Status do sistema exibido corretamente | Status do sistema exibido corretamente | Status do sistema exibido corretamente | Status do sistema exibido corretamente |
| Combinação entre o sistema e o mundo real | Não, incorreta e confusa | Não retornou resposta e não informa motivo | Não sabe/Não reconhece cancelamento de médico | Compreendeu corretamente | Compreendeu corretamente |
| Liberdade e controle do usuário | Script não exige ação para esta heurística | Script não exige ação para esta heurística | Não dá a liberdade de cancelar o médico | Script não exige ação para esta heurística | Não permite pular ou avançar nas respostas |
| Consistência e padrões | Script não exige ação para esta heurística | Script não exige ação para esta heurística | Script não exige ação para esta heurística | Script não exige ação para esta heurística | Script não exige ação para esta heurística |
| Prevenção de erro | Script não exige ação para esta heurística | Script não exige ação para esta heurística | Não conseguiu cancelar o evento | Script não exige ação para esta heurística | Script não exige ação para esta heurística |
| Reconhecimento em vez de recordação | Script não exige ação para esta heurística | Script não exige ação para esta heurística | Resposta simples e fácil de lembrar | Resposta simples e fácil de lembrar | Péssimo para lembrar de todas as informações |
| Flexibilidade e eficiência de uso | Difícil saber o que deu errado | Difícil saber o que deu errado | Difícil saber o que deu errado | Resposta simples e satisfatória | Não é eficiente para pular respostas |
| Design estético e minimalista | Não há clareza na resposta | Não há clareza na resposta | Não há clareza na resposta | Resposta simples e fácil de lembrar | Possui informações em excesso e não há garantia que estejam corretas |
| Ajude os usuários a reconhecer, diagnosticar e recuperar erros | Não orientou nem solucionou o problema | Não orientou nem solucionou o problema | Não orientou nem solucionou o problema | Orientou como resolver o problema | Não soluciona nem orienta como voltar ao início |
| Ajuda e documentação | Não fornece ajuda nem documentação | Não fornece ajuda nem documentação | Não fornece ajuda nem documentação | Script não exige ação para esta heurística | Script não exige ação para esta heurística |

Análise dos resultados obtidos a partir da execução dos testes

Após a execução dos testes, foram feitas as análises dos resultados obtidos. As tabelas geradas como descritas na seção anterior, facilitam as verificações e as conclusões. Embora sejam conclusões com base na percepção da autora como uma usuária comum, os resultados obtidos foram semelhantes aos obtidos por Nielsen (que serão descritos logo adiante).

A interface de voz Alexa apresenta uma boa interação para consultas simples, tais como temperatura, alarme e notícias. As respostas são diretas e fáceis de entender. Porém, fazer questões que podem ter um duplo sentido, usar um dialeto comum como “cachorro de porte médio”, ou solicitar apenas hotéis em um determinado bairro (que não garantem a veracidade da resposta), além de prolongar uma conversa com ela com “próxima”, “voltar”, tiveram um desempenho ruim quando comparado às questões curtas sem outro sentido.

A interface apresenta muito bem a heurística 1, de visibilidade de status de sistema; mostra corretamente um círculo azul onde deve ser clicado para se fazer a pergunta e uma sinalização de quando está escutando por uma linha na parte inferior do aplicativo; a interface se saiu bem em todas as perguntas.

Na heurística 2, relativa ao mundo real, a interface se sai bem para questões simples e curtas, como informar temperatura, pesquisar notícia ou programar alarme. Para questões que envolvam duas sentenças ou frases que possam ter um sentido duplo, a interface apresenta problemas de reconhecimento ou, em alguns casos, apresenta respostas confusas, inclusive em inglês (questões 5 a 10).

Para a heurística 3, de controle do usuário, ela teve um desempenho bom quando se pediu para parar de falar, mas ruim quando foi solicitado o cancelamento do alarme, do médico e para ir para o próximo item da lista no meio da leitura.

Para a heurística 4, de consistência e padrões, poucas questões se adequaram a elas, pois as heurísticas possuem características que as questões propostas não tinham para testar (no caso, apenas uma foi falada de maneira semelhante, a solicitação de pesquisa sobre cachorro, para a qual a interface não deu a resposta requerida).

Para a heurística 5, de prevenção de erro, não houve questões que pudessem ser aplicadas a ela, pois apenas a do alarme teve uma confirmação implícita de maneira correta.

Quanto à heurística 6 (recordação), a interface deu boas repostas para questões curtas e diretas. Para a questão 10 deu uma resposta longa, lendo todas as seis opções que encontrou, sendo impossível se lembrar de todas.

Na heurística 7 (flexibilidade e eficiência de uso), a interface mostra que não é tão simples usá-la. Um usuário leigo ou mais experiente ficaria em dúvida sobre como usá-la sem precisar escutar tudo o que ela fala, ou não saberia como resolver os problemas das respostas incorretas.

Na heurística 8 (design estético e minimalista), houve uma boa interação nas cinco primeiras questões, porém nas questões seguintes começou a apresentar respostas pouco claras sobre o que devia ser feito. A última questão apresentou a pior avaliação quanto a esta heurística, pois as informações foram demasiadas tanto na tela quanto na fala.

Nas heurísticas 9 e 10 (ajuda e recuperação de erro e documentação), a interface também não satisfaz as expectativas. Para a maioria das questões que ela não conseguiu ou não soube dar a resposta, não orientou como deveria ser perguntado nem ofereceu ajuda verbal ou visual.

Percebeu-se também que é necessário pensar como fazer a pergunta, ou repeti-la quando se dá uma pausa na frase. Isto faz o uso se tornar cansativo. Outro problema foi a fala. Além de a Alexa falar muito rápido, é quase impossível, dependendo da resposta, memorizar tudo o que ela diz (isto atinge a heurística minimalista).

Foi necessário fazer tentativa e erro até encontrar as respostas, ou desistir de usá-la e refazer a pesquisa em um navegador. Isto afeta imensamente a questão da usabilidade, que está relacionada com a facilidade de aprendizado e uso da interface, bem como a satisfação do usuário em decorrência desse uso (Verificar Conceitos).

Outra questão é a qualidade da resposta dada em tela, baseada em uma pesquisa em que não se garante a veracidade do que foi respondido (como na questão da temperatura e dos hotéis).

Outro ponto foi o controle do que se ouvia. Não há como pular para próximo, voltar ou reler. É quase obrigatório ficar escutando-a falar e se perder em meio a muitas informações.

Budiu (2018) comenta que Nielsen também propôs vários testes com interfaces como Siri, Google e Alexa com participantes e o resultado foi semelhante. Interação limitada e dificuldade em entender como utilizar ou o que se deve dizer. Explicou em um dos seus testes: “Nossa pesquisa de usuário descobriu que os assistentes inteligentes atuais falham em todas as 6 perguntas (5 tecnologias mais integração), resultando em um nível de usabilidade geral que é quase inútil, mesmo para interações ligeiramente complexas. Para interações simples, os dispositivos atendem aos requisitos mínimos de usabilidade. Mesmo que vá contra a premissa básica do design centrado no ser humano, os usuários devem treinar para entender quando um assistente inteligente será útil e quando é melhor evitar usá-lo”.

Como Nielsen mesmo diz: Os assistentes "inteligentes" de hoje ainda estão longe de passar no teste de Turing. Para a maioria das interações, as pessoas descobrirão facilmente que não estão falando com um humano. Embora os usuários projetem qualidades humanas neles, eles têm expectativas relativamente baixas em relação a esses assistentes e os reservam para questões reais e em preto-e-branco. Mesmo que o principal obstáculo seja provavelmente a linguagem natural e o processamento de diálogo (um problema inerentemente difícil), muitos problemas de menor escala poderiam ser corrigidos com um design mais cuidadoso (BUDIU, 2018).

CONCLUSÃO E POSSÍVEIS SOLUÇÕES DOS TESTES

As interfaces de voz ainda precisam evoluir bastante. Claro que se trata de algo programado pelo ser humano que tem suas qualidades e defeitos, porém elas ainda não atingiram o ponto de parecer que realmente estamos falando (tão naturalmente) com outro ser humano. A interface não é ilimitada e está longe disso. Porém notam-se frustrações ao usá-la para determinadas funções, em específico para manter uma conversa e controlar o que é dito e fazer perguntas mais difíceis. Neste caso, proponho algumas possíveis soluções para melhorias das interfaces de voz:

1-Melhorar a compreensão da entrada e saída de voz (há problemas com frases ditas muito devagar quando se faz uma pausa ao longo da questão);

2-Melhorar a interação da voz. A interface fala muito rápido para questões em que se deseja gravar alguma coisa, e é difícil se recordar de tudo;

3-Tentar dar o mínimo de respostas na tela. Embora seja eficiente quando alguém deseja anotar as várias opções, ela exibe muitas opções na tela e ao mesmo tempo as lê. As interfaces de voz, como diz Nielsen, são projetadas para a interação verbal e ajudar aqueles que estão com mãos ocupadas ou não desejam usá-las (as mãos);

4-Orientar e prevenir erros. Um usuário precisa testar e testar novamente. Isto atinge o conceito de usabilidade que envolve a facilidade de uso. Perguntar coisas como “você pode tentar dizer...” e “caso desejar saber sobre tal assunto navegue” facilita o uso e ajuda a prevenir erros;

5-Oferecer uma documentação simples e fácil. É necessário o usuário pesquisar no Google ou adivinhar como usar ou dizer algo a ela. Oferecer um manual de voz quando verificado um erro ou disponibilizar no aplicativo auxiliaria, caso o mesmo precise;

6-Permitir que o usuário tenha controle sobre o que ouvir e permitir que pule, volte ou escute novamente o que deseja;

7-Mostrar ao usuário o que está errado. Demonstrar que não conhece a palavra, frase ou parte dela e como o usuário deve proceder;

8-Evitar que o usuário tente fazer sempre as mesmas questões;

9-Compreender questões encadeadas e de duplo sentido, ou, quando isto ocorrer, saber como perguntar quais dos dois sentidos deve ser considerado.

Embora algumas soluções afetem a questão de usabilidade, oferecer algo com que o usuário possa aprender satisfaria suas necessidades, até que esses dispositivos se tornem inteligentes o suficiente para não precisarem mais disso.

CONCLUSÃO

O artigo mostrou, ao longo das seções, que as interfaces evoluíram drasticamente ao longo dos anos. Desde os anos 1950, havia empresas interessadas em desenvolver uma máquina/sistema que pudesse compreender e interagir com o ser humano. Assim, os sistemas ficaram cada vez mais modernos e ao longo dos anos surgiram a Siri, Alexa, Cortana, Google Assistant, entre outros, atingindo uma enorme popularidade. A Alexa, estando em muitos dispositivos inteligentes, se tornou muito comum.

Assim também aumentaram as pesquisas na área de inteligência artificial, para que estes dispositivos se tornassem cada vez melhores. Questões como usabilidade, ergonomia, acessibilidade e IHC se tornaram de grande valia para o desenvolvimento destas interfaces. A usabilidade, acessibilidade e aceitabilidade permitem que sistemas sejam construídos para atender os usuários da melhor forma possível.

O estudo das 10 heurísticas de Nielsen demonstra o quanto elas são importantes para o desenvolvimento de interfaces, e Mittal descreveu como podem ser validadas em interfaces de voz. Concluiu-se que estudos de interfaces visuais podem ser aplicados para voz, porém devem ser adaptados para o contexto e nem sempre são eficazes.

O teste no qual a autora verifica como a Alexa se comporta mostrou resultados variados. A interface se saiu bem com questões simples e não teve boa interação nas questões mais longas, além de não tratar os erros dados, deixando a conversa confusa e o usuário sem saber o que fazer.

Nem toda questão foi apta a ser aplicada a determinada heurística, pois não exibiu uma ação que a heurística exigia. Mesmo assim, do ponto de vista da autora como usuária, a Alexa deixa a desejar em muitos aspectos, limitando-se a tarefas simples; dependendo da situação, seu uso é ineficaz, sendo melhor usar o navegador de internet.

O teste realizado pela autora se assemelhou ao que Nielsen afirmou sobre os testes que ele fez com usuários, revelando as limitações das interfaces de voz e que elas precisam melhorar muito para atingir as

exigências de todos os usuários. Por fim, algumas possíveis soluções foram apresentadas, com base na visão da autora para a melhoria da Alexa.

Concluimos que as interfaces estão longe de ser perfeitas e que ainda não podemos exigir delas o quanto gostaríamos, bastando nos adequar ao que são capazes no momento e aguardar as suas promissoras evoluções no futuro.

REFERÊNCIAS BIBLIOGRÁFICAS

- LEIS de Clarke: É Tecnologia Avançada indistinguível de magia? **GRELLANE**, 2017. Disponível em: <<https://www.grellane.com/pt/ci%C3%A4ncia-tecnologia-matem%C3%A1tica/ci%C3%A4ncia/what-are-clarkes-laws-2699067/>>. Acesso em: 27 de nov. de 2020.
- COATES A., Dustin. Voice Applications for Alexa and Google Assistant. NY.USA: Manning Publications Co. 2019.
- QUESADA, W., LAUTENBACH, B., Programing Voice Interfaces: Giving Connected Devices a Voice. CA, USA, O'Reilly Media. 2018.
- PEARL, Cathy. Design Voice User Interface: Principles of Conversational Experiences. CA, USA, O'Reilly Media. 2017.
- ALLIANNEJADI, M, RÍSSOLA, E A, CHAKRABORTY, M, CRESTANI, F. **Harnessing Evolution of Multi-Turn Conversations for Effective Answer Retrieval**, CHIIR '20, Vancouver, 2020.
- SCIUTO, A, SAINI, A, FORLIZZI, J, HONG, J I. **“Hey Alexa, What’s Up?”: Studies of In-Home Conversational Agent Usage**, Carnegie Mellon University, Pittsburgh, 2018. (HONG KONG)
- SIEGERT, I, KRUGER, J. How do we speak with ALEXA – Subjective and objective assessments of changes in speaking style between HC and HH conversations, University Magdeburg, 2018. (Germany)
- BARBOSA J., S D., SILVA, B. Interação Humano-Computador. Rio de Janeiro: Elsevier, 2010.
- SILVA G, Kim. Assistentes de voz presentes em alto-falantes inteligentes: Uma análise exploratória sobre os tópicos de pesquisa e as possibilidades de uso. Escola de Comunicação Artes e Design Famedcos. 2019 (Porto Alegre)
- CRUZ T, L., ALENCAR, A J, SCHMITZ, E A. Assistentes Virtuais Inteligentes: Conceitos e Estratégias. Rio de Janeiro: Brasport, 2013.
- FIORETTI, Livia. Por que a tecnologia tem voz de mulher (e por que isso importa?). **O futuro das coisas**, 2018. Disponível em: <<https://ofuturodascoisas.com/porque-a-tecnologia-tem-voz-de-mulher-e-porque-isso-importa/>>. Acesso em: 27 de nov. de 2020.
- STOIDER L., R. Steve Jobs reprovava o nome “Siri”, diz executivo. **Adrenaline**, 2012. Disponível em: <<https://adrenaline.com.br/noticias/v/11190/steve-jobs-reprovava-o-nome-siri-diz-executivo>>. Acesso em: 27 de nov. de 2020.
- MACEDO M., G. 10 Heurísticas de Nielsen para o Design de Interface. **UX Collective**, 2017. Disponível em: <<https://brasil.uxdesign.cc/10-heur%C3%ADsticas-de-nielsen-para-o-design-de-interface-58d782821840>>. Acesso em: 27 de nov. de 2020.
- MITTAL, D. How Nielsen’s 10 usability heuristics apply to Voice UI. **UX Collective**, 2018. Disponível em: <<https://uxdesign.cc/10-usability-heuristics-for-voice-user-interface-design-69ad9ea4f166/>>. Acesso em: 27 de nov. de 2020.

BUDIU, R. **Intelligent Assistants Have Poor Usability: A User Study of Alexa, Google Assistant, and Siri.** Nielsen Norman Group, 2018. Disponível em: <<https://www.nngroup.com/articles/intelligent-assistant-usability/>>. Acesso em: 27 de nov. de 2020.

PEREIRA, S L. Processamento de Linguagem Natural. USP, 1997.

WHITENTON, k., BUDIU, R. **The Paradox of Intelligent Assistants: Poor Usability, High Adoption.** Nielsen Norman Group. 2018. Disponível em: < <https://www.nngroup.com/articles/intelligent-assistants-poor-usability-high-adoption/>>. Acesso em: 27 de nov. de 2020.