

Projetando para a voz no contexto bancário: uma proposta de interface multimodal para o aplicativo do Banco do Brasil

Diseñando para la voz en el contexto bancario: una propuesta de interfaz multimodal para la aplicación del Banco do Brasil

Designing for voice in the banking context: a multimodal interface proposal for the Banco do Brasil application

Maurício Elias Klafke Dick, Anderson Massao Tanaka

design digital, interface por voz do usuário, projeto, aplicativo, Banco do Brasil

Diante da popularização dos assistentes de voz no setor financeiro, este estudo investiga as limitações da interface atual do aplicativo do Banco do Brasil quanto ao uso de comandos de voz e propõe uma forma de interação natural por meio de uma interface multimodal baseada em comandos de voz. A metodologia adotada foi baseada no modelo de Garrett (2011), que estrutura o projeto em cinco planos: Estratégia, Escopo, Estrutura, Esqueleto e Superfície. Como resultado, tem-se a proposta de um sistema integrado à interface gráfica do aplicativo, que busca proporcionar uma experiência bancária mais inclusiva, prática e acessível.

diseño digital, interfaz mediante voz del usuario, proyecto, aplicación, Banco do Brasil

Dada la popularización de los asistentes de voz en el sector financiero, este estudio investiga las limitaciones de la interfaz actual de la aplicación del Banco do Brasil con respecto al uso de comandos de voz y propone una forma de interacción natural a través de una interfaz multimodal basada en comandos de voz. La metodología adoptada se basó en el modelo de Garrett (2011), que estructura el proyecto en cinco planes: Estrategia, Alcance, Estructura, Esqueleto y Superficie. Como resultado, proponemos un sistema integrado en la interfaz gráfica de la aplicación, que busca brindar una experiencia bancaria más inclusiva, práctica y accesible.

digital design, voice user interface, project, application, Banco do Brasil

Given the popularization of voice assistants in the financial sector, this study investigates the limitations of current Banco do Brasil application interface regarding the use of voice commands and proposes a way of natural interaction through a multimodal interface based on voice commands. The methodology was based on the Garrett model (2011), which structures the project into five planes: Strategy, Scope, Structure, Skeleton and Surface. As result, we have proposed a system integrated into the application's graphical interface, which seeks to provide a more inclusive, practical and accessible banking experience.

1 Introdução

O ser humano absorve e compreende informações em diferentes formas através dos seus sentidos, seja como palavras, imagens, movimento ou som; e, na perspectiva do design, busca-se torná-las significativas ao assegurar a sua comunicação efetiva (Baer, 2021). Assim,

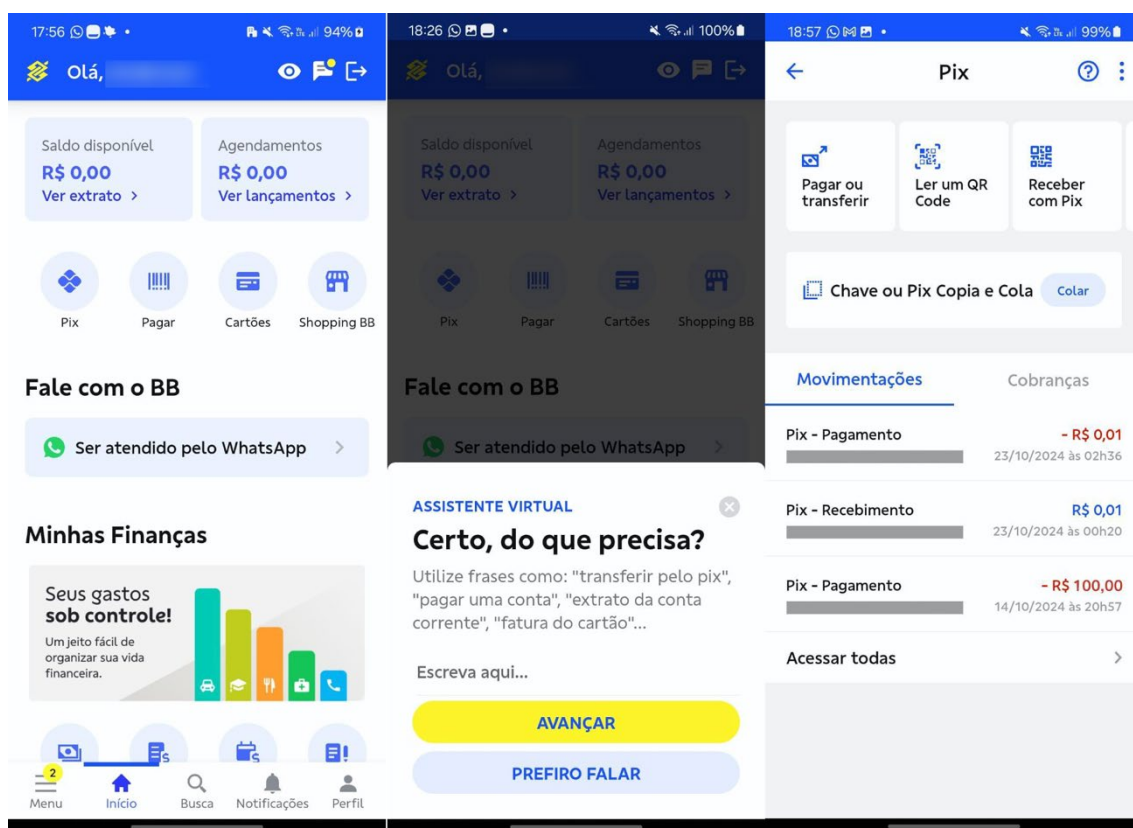
em razão dessa essência, o design da informação é uma área interdisciplinar e centrada no ser humano e, no contexto digital, converge com outros campos como arquitetura da informação, design de interação, design de interfaces e design para a experiência do usuário (Baer, 2021). Um dos pilares para desenvolver uma interação que cumpra com seu objetivo é compreender as necessidades e expectativas do usuário, pois, a partir dessas informações, é possível criar soluções que atendam às demandas de forma direta e intuitiva. Como apontado por Sharp *et al.* (2019), entender o comportamento das pessoas e a natureza das interações com sistemas tecnológicos é fundamental para o desenvolvimento de interfaces plenamente funcionais.

Ao observar o comportamento humano na interação com as tecnologias bancárias, segundo a Federação Brasileira de Bancos (Deloitte Global, 2024a), houve um aumento de 22% nas transações realizadas por dispositivos móveis entre 2022 e 2023, seguindo movimentos de expansão anteriores, indicando uma mudança nos hábitos dos clientes em direção a um ambiente bancário mais digitalizado. Entre as prioridades estratégicas dos bancos em 2024 (Deloitte Global, 2024b), destaca-se a expansão de transações mediadas por *chatbots*. Além disso, a inteligência artificial (IA) tem sido considerada uma prioridade para aprimorar a experiência do cliente, garantindo maior grau de personalização (Deloitte Global, 2024b).

Essas mudanças no cenário bancário, como a integração de IA e o uso crescente de *chatbots* para interações e operações financeiras, têm criado um ambiente favorável para o desenvolvimento das interfaces por voz (*Voice User Interface* – VUI). Inseridas no contexto das interfaces naturais (*Natural User Interface* – NUI), essas interfaces operam por comandos ou conversas em que os usuários ouvem e falam com um sistema, em vez de clicar, tocar ou apontar para uma tela (Sharp *et al.*, 2019). Elas ganharam grande popularidade nos últimos anos, principalmente com o advento de assistentes de voz como Siri, Alexa e Google Assistente. Com a evolução tecnológica, este tipo de interação utiliza processamento de linguagem natural (NLP) e inteligência artificial para interpretar comandos verbais e responder ao usuário de forma personalizada (Sharp *et al.*, 2019).

Segundo pesquisa da revista Exame (2024), o Banco do Brasil é o segundo maior banco do país, mas figura como o primeiro dentro dos bancos públicos, com patrimônio ativo de aproximadamente 2,3 trilhões de reais. Comparado a outros grandes bancos, como Caixa Econômica Federal, Bradesco, Itaú e Nubank, o Banco do Brasil se destaca como um banco público de ampla capilaridade e já oferece algum nível de interação baseada em comandos naturais, embora ainda limitada em funcionalidades e acessibilidade. A interface por comando de voz existente no aplicativo do Banco do Brasil, atualmente, não oferece suporte completo a operações por meio da voz e apenas direciona o usuário para a tela da função requisitada, como mostra a Figura 1.

Figura 1: Exemplo de uso da assistente virtual do aplicativo BB para sistemas Android. Fonte: Banco do Brasil SA (2024).



Por conta dessa limitação, identificou-se a oportunidade de desenvolvimento de uma *Voice User Interface* (VUI) mais robusta para o Banco do Brasil, capaz de auxiliar o usuário em operações bancárias utilizando comandos de voz. Segundo Doherty e Curran (2019), essas interfaces representam uma tendência em expansão e é uma resposta à demanda dos indivíduos por rapidez e simplicidade no acesso a informações bancárias, e sua aplicação nos aplicativos de bancos pode oferecer uma experiência de usuário mais acessível e inovadora. Em relação a outras tecnologias de autoatendimento, como as centrais telefônicas existentes há diversos anos, o uso de linguagem natural e inteligência artificial possibilita uma experiência mais fluida e autônoma, sem menus hierárquicos rígidos e a necessidade de operadores humanos.

Assim, esta pesquisa teve como objetivo propor uma forma de interação natural com o aplicativo do Banco do Brasil, por meio de uma interface multimodal baseada em comandos de voz. Nesse sentido, busca-se criar um sistema que faça uso da linguagem auditiva apoiada pela linguagem visual gráfica, explorando os modos verbal e pictórico de comunicação (Twyman, 1982). Para isso, adota-se a abordagem do Design Science Research (Santos, 2018) e desenvolve-se a interface com base no modelo de Garrett (2011), que organiza o projeto para experiência do usuário em cinco planos: Estratégia, Escopo, Estrutura, Esqueleto e Superfície.

2 Interação natural e as particularidades da interface por voz

Existem diferentes tipos de interfaces que facilitam a interação do usuário com os sistemas. Dentre elas, tem-se as interfaces naturais (*Natural User Interface* – NUI), que são projetadas para permitir que os usuários interajam com os sistemas utilizando habilidades intrínsecas do ser humano, como movimentos do corpo, fala ou toques (Sharp *et al.*, 2019). As NUIs procuram integrar a tecnologia ao contexto da pessoa, proporcionando uma experiência que parece menos "digital" e mais próxima das interações existentes no mundo real, minimizando a curva de aprendizado e permitindo que o usuário interaja com o sistema de forma fluida.

Inseridas nesse contexto, as interfaces por voz do usuário (*Voice User Interface* – VUI) são um modelo de interação humano-computador que tem como base o uso da voz para a realização de tarefas (Dasgupta, 2018). Desse modo, elas possibilitam uma interação natural e promovem maior acessibilidade e inclusão, oferecendo alternativas eficazes para indivíduos que possam ter limitações físicas ou cognitivas, além de simplificar a experiência para todos os usuários (Sharp *et al.*, 2019). Em comparação a outros modos de interação – como gestos ou toques –, as interfaces por voz do usuário oferecem interação sem o uso das mãos e dos olhos, permitem que os usuários executem outras ações simultaneamente, são análogas às conversas cotidianas e são benéficas em termos de acessibilidade (Iliev & Ilieva, 2023).

Caracterizada por ser linear no tempo, unidirecional e invisível (Bouazid & Ma, 2022), esta tecnologia é estudada já há diversos anos e tem se tornado amplamente disponível mais recentemente (Dasgupta, 2018), uma vez que melhorias significativas passaram a permitir que as pessoas se comuniquem com os dispositivos de uma forma mais natural e intuitiva (Klein, 2015). A era dos assistentes digitais simboliza um novo momento das VUIs (Dasgupta, 2018) e marcam um período de interação independente de telas (Bouazid & Ma, 2022). Com isso, essa tecnologia tem se tornado cada vez mais presente no dia a dia das pessoas, refletindo uma mudança significativa na maneira como elas interagem com os dispositivos computacionais.

Projetar interfaces por voz guarda similaridades com o design de outros tipos de interface, em termos de princípios e métodos (Klein, 2015; Cohen, Giangola & Balogh, 2004). As pessoas devem ser o centro das decisões de projeto, portanto, compreender o usuário, conhecer o contexto de uso e garantir que os produtos sejam úteis e utilizáveis são questões que se mantêm pertinentes (Klein, 2015). Por outro lado, privacidade e acessibilidade se tornam preocupações que demandam ainda mais atenção (Klein, 2015). Além disso, ao projetar para voz, os designers precisam ser cuidadosos para não sobrecarregar os usuários, mantendo as conversações simples, curtas e eficazes (Dasgupta, 2018). Nesse sentido, interfaces que combinam a fala com outras modalidades de interação (interfaces multimodais) podem ser uma estratégia para projetar interações que reduzam a demanda cognitiva do usuário (Cohen, Giangola & Balogh, 2004). Para Dasgupta (2018), é importante que um sistema baseado em voz também seja capaz de permitir interações por meio de diferentes modos, tais como toque, gestos e informações visuais, por exemplo. Nessa direção, as interfaces multimodais podem aumentar a facilidade de uso, permitindo uma interação mais natural na realização de tarefas, e podem acomodar uma ampla gama de usuários, incluindo pessoas com deficiências

temporárias ou permanentes e outros casos em que uma modalidade pode não ser suficiente (Neto *et al.*, 2008).

No que tange ao projeto de interfaces por voz, suas entregas geralmente incluem exemplos de diálogos, diagramas de fluxo, listas de *prompts* e simulações das telas (no caso das interfaces multimodais) (Pearl, 2016). Cabe destacar que os diagramas de fluxo, ou fluxogramas, ilustram os principais caminhos que podem ser percorridos pelo usuário por meio da interface (Pearl, 2016), enquanto, no contexto das VUIs, *prompts* são as mensagens do sistema (Cohen, Giangola & Balogh, 2004). Outro ponto importante que não deve ser negligenciado pelos designers é a personalidade do sistema, isto é, sua imagem perante os usuários, que compreende desde o tom de voz até o uso de diferentes tipos de avatar e suas expressões para representar o *bot* na interação humano-máquina (Dasgupta, 2018; Cohen, Giangola & Balogh, 2004).

3 Metodologia

Esta pesquisa se desenvolve a partir da abordagem do Design Science Research (Santos, 2018), na medida em que a prática projetual foi utilizada como meio de investigação e geração de conhecimento. A adoção dessa estratégia metodológica é adequada “quando há a criação de um artefato para a promoção de melhorias no mundo real presente ou futuro” (Santos, 2018, p. 76). Para tal, a metodologia utilizada para o desenvolvimento da pesquisa é baseada no modelo de Jesse James Garrett (2011), que propõe um olhar sobre o projeto de interfaces segmentado em cinco planos: Estratégia, Escopo, Estrutura, Esqueleto e Superfície.

Conforme a visão do autor (Garrett, 2011), foram estabelecidos os seguintes objetivos e procedimentos para cada plano:

- **Estratégia:** Especificar as necessidades e as expectativas dos futuros usuários em relação ao uso de assistentes de voz em bancos. Para tal, utilizaram-se os dados secundários da pesquisa FEBRABAN de Tecnologia Bancária¹. A análise do conteúdo destes relatórios foi realizada com o uso de ferramenta de inteligência artificial, cujos *insights* permitiram elaborar proto-personas;
- **Escopo:** Definir as funcionalidades da interface de voz para o Banco do Brasil. Com este objetivo, foi realizada uma análise de similares e novamente usados dados secundários da pesquisa FEBRABAN de Tecnologia Bancária. Na sequência, utilizou-se uma técnica de priorização de requisitos que permitiu organizá-los em obrigatórios, desejáveis e opcionais;
- **Estrutura:** Planejar o fluxo de navegação e as interações possíveis entre o usuário e a interface de voz, priorizando um design que fosse fácil e intuitivo para uso. Para isso,

¹ A coleta de dados é feita por meio de formulários eletrônicos e entrevistas em profundidade, complementada por dados públicos e de pesquisas anteriores. Em suas últimas edições, foram representados, pelo menos, 80% dos ativos bancários do Brasil (Deloitte Global, 2024a, 2024b, 2023a, 2023b).

foi elaborado um diagrama de arquitetura da informação e realizada a decomposição das tarefas por meio de diagramas de fluxo (fluxogramas);

- **Esqueleto:** Desenvolver *wireframes* e protótipos que representassem e apoiassem a interação por voz e seus principais componentes. Com este intuito, foram realizadas análise de referências, geração de alternativas e prototipação;
- **Superfície:** Manter uma coerência visual com o aplicativo do Banco do Brasil, assegurando uma integração com a interface gráfica do usuário. Para tal, foram analisados o Manual de Identidade Visual da instituição e as telas do aplicativo “BB: Banco, Conta, Pix, Crédito”.

4 Resultados

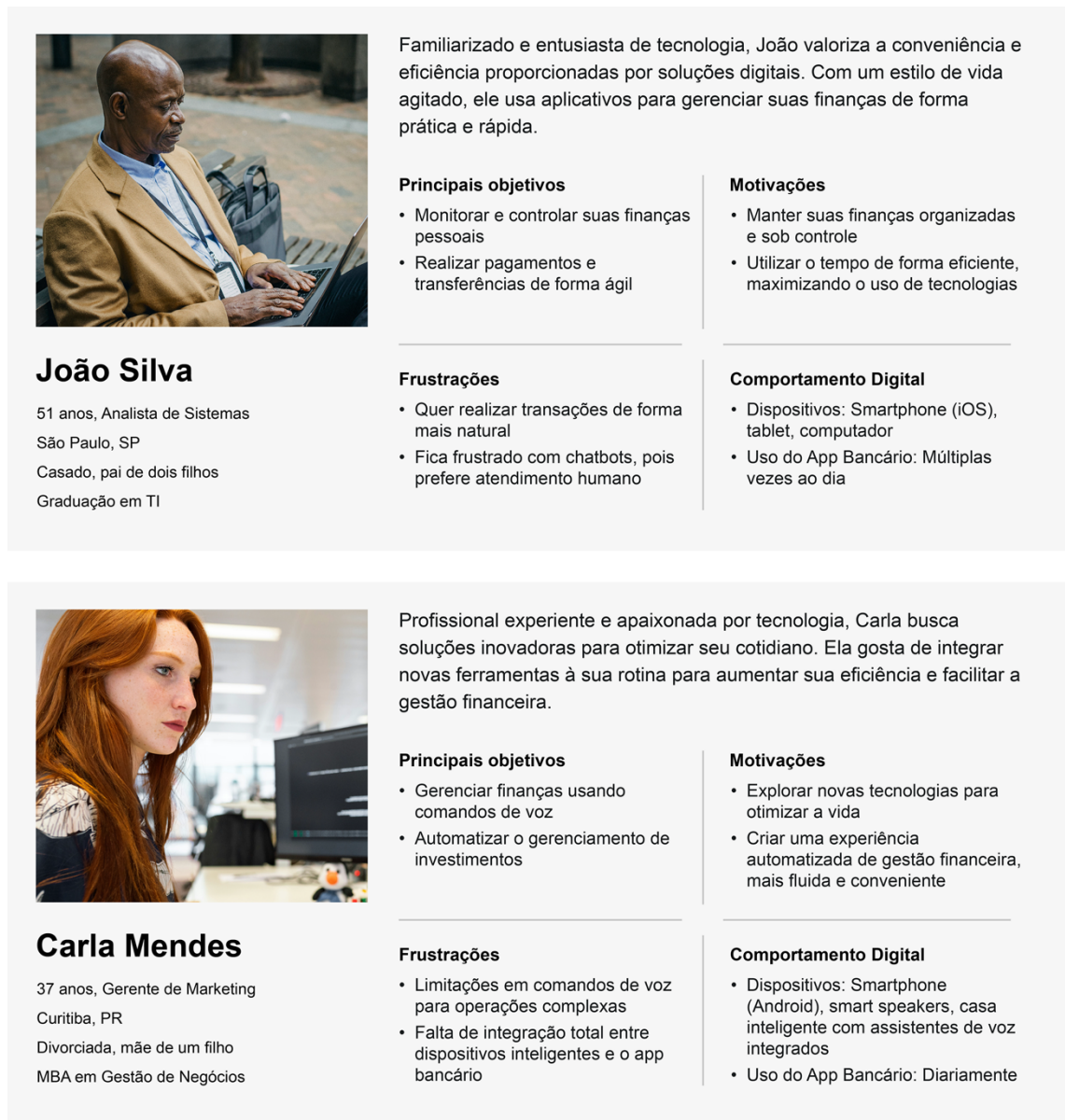
4.1 Estratégia

Para especificar as necessidades e as expectativas dos futuros usuários em relação ao uso de assistentes de voz em bancos, foram desenvolvidas proto-personas utilizando dados da pesquisa FEBRABAN de Tecnologia Bancária 2023 (Deloitte Global, 2023a; 2023b). As proto-personas são representações fictícias de um grupo específico de usuários ou clientes potenciais de um negócio ou produto, sem que haja pesquisa de campo ou observação dos usuários (Tomlin, 2018).

Em síntese, os dados indicam que os clientes dos bancos brasileiros buscam soluções rápidas, fáceis e convenientes, e esperam um atendimento cada vez mais digital e humanizado. Eles desejam realizar diferentes tipos de transações bancárias em qualquer lugar e consultar sua situação financeira pessoal a qualquer momento, de forma simples e fluida. No entanto, existem frustrações, como a baixa usabilidade e pouca integração dos recursos de autoatendimento digital.

Em termos de comportamento digital, os clientes demonstram uma alta adoção ao *mobile banking*, sendo o canal preferido para a maioria das transações. Quase 8 em cada 10 transações bancárias são realizadas via canais digitais e 48% dos usuários cadastrados no sistema Pix são *heavy users*, isto é, realizaram mais de 30 transações deste tipo por mês. A partir da extrapolação destas constatações, foi possível criar duas proto-personas (Figura 2).

Figura 2: Proto-personas do projeto. Fonte: autores (2025).



Ambas as proto-personas revelam necessidades comuns no uso de soluções financeiras digitais. Os dois usuários desejam simplificar e otimizar o gerenciamento de suas finanças de forma prática, valorizando agilidade e eficiência nas operações bancárias. Entre suas principais frustrações está a falta de integração com comandos de voz para operações mais avançadas. Por fim, suas motivações giram em torno de maximizar o uso de tecnologias para gerenciar suas finanças de maneira fluida e controlada, mantendo a rotina organizada.

Com base nestas inferências, estabeleceu-se que o sistema teria como objetivo simplificar processos e otimizar operações financeiras, promovendo maior inclusão por meio de uma interface natural que utilizasse a fala como principal meio de interação.

4.2 Escopo

Com o objetivo de definir as funcionalidades da interface de voz para o Banco do Brasil, inicialmente foi realizada uma análise de sistemas similares, observando as interfaces conversacionais dos principais concorrentes da instituição à época de realização da pesquisa (outubro de 2024). Foram examinadas as assistentes virtuais dos bancos Caixa Econômica Federal, Itaú, Bradesco, Santander, Nubank e Inter, conforme ilustra a Figura 3.

Figura 3: Análise de similares. Fonte: compilado dos autores a partir de Caixa Econômica Federal (2024), Itaú Unibanco Holding S.A. (2024), Banco Bradesco SA (2024a), Banco Santander (Brasil) S.A. (2024), Nu (2024) e Banco Inter SA (2024).

	Caixa Econômica Federal	Itaú	Bradesco	Santander	Nubank	Inter
Chatbot	✓	✓	✓	✓	✓	✓
Realiza operações	✗	?	✓	✗	✓	✗

Apenas dois² dos seis bancos analisados possuem uma interface multimodal desenvolvida (Bradesco³ e Nubank), permitindo realizar operações bancárias por meio de comandos de voz. Já os outros quatro bancos se restringem a *chatbots* para tirar dúvidas frequentes ou específicas dos usuários e direcioná-los a seções dos aplicativos, não possuindo uma interface natural que permita realizar transações por meio da voz. É importante notar que em todos os sistemas analisados, o usuário precisa digitar o que deseja e a interface responde por texto escrito em um chat. Em três destes (Bradesco, Nubank e Itaú), porém, é possível ao usuário falar o que deseja e tal *input* é transcrito na conversa.

Para auxiliar na definição das principais funcionalidades da interface a ser desenvolvida, utilizou-se a pesquisa FEBRABAN de Tecnologia Bancária 2024. Segundo o estudo (Deloitte Global, 2024b), em 2023 as operações financeiras mais realizadas por meio de dispositivos móveis são (em milhões de transações): saldos e extratos (73.938), Pix (16.796), pagamento de contas (3.051), consultas de cartão de crédito (1.283), contratação de crédito (1.283), transferência/DOC/TED (818) e consulta de investimentos (288), consolidando uma gama de

² Não foi possível ter acesso à assistente virtual do Banco Itaú para avaliar se a plataforma permite a realização de operações. Em seu site (Itaú Unibanco Holding S.A., 2024), a instituição afirma que é possível consultar saldo, ajustar limites, simular empréstimos, renegociar dívidas, tirar dúvidas, entre outros.

³ Também não foi possível ter acesso à assistente virtual do Banco Bradesco. Entretanto, conforme o banco, “no app Bradesco, a BIA te informa sobre o saldo da conta e [...] também faz transações, é só falar ou digitar o que quer fazer” (Banco Bradesco, 2024b).

funcionalidades que são frequentemente acessadas remotamente pelos usuários. Na sequência, estabeleceram-se os requisitos para o projeto, categorizados em obrigatórios, desejáveis e opcionais (Quadro 1), com base nos dados analisados até então (Estratégia e Escopo).

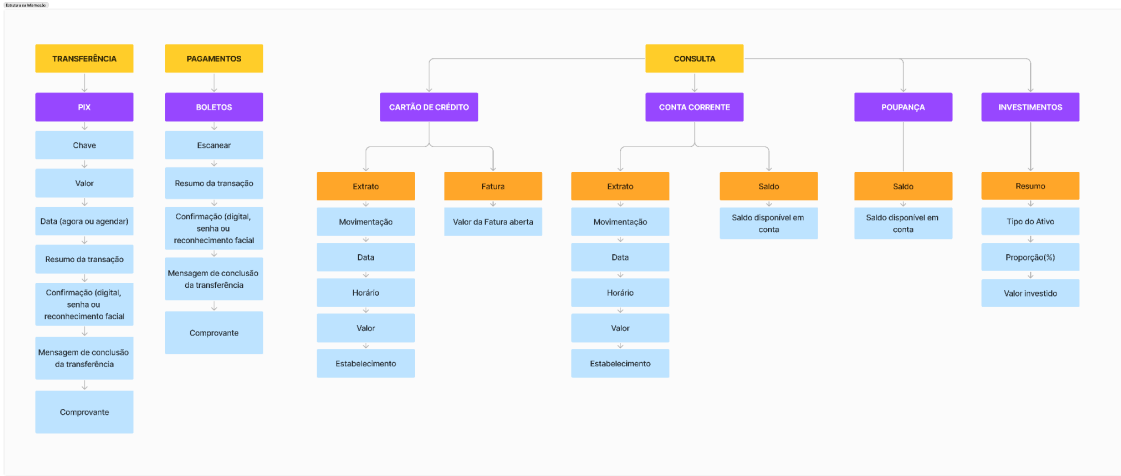
Quadro 1: Hierarquia dos requisitos do projeto.

Prioridade	Requisitos
Obrigatórios	Transferências e Pagamentos Consulta de Saldo e Extrato Consulta de Investimentos Autenticação Segura Comprovante de Transação Feedback
Desejáveis	Personalização da Interface Notificações Inteligentes Interface Conversacional Avançada Agendar Operação Financeira
Opcionais	Saúde Financeira

4.3 Estrutura

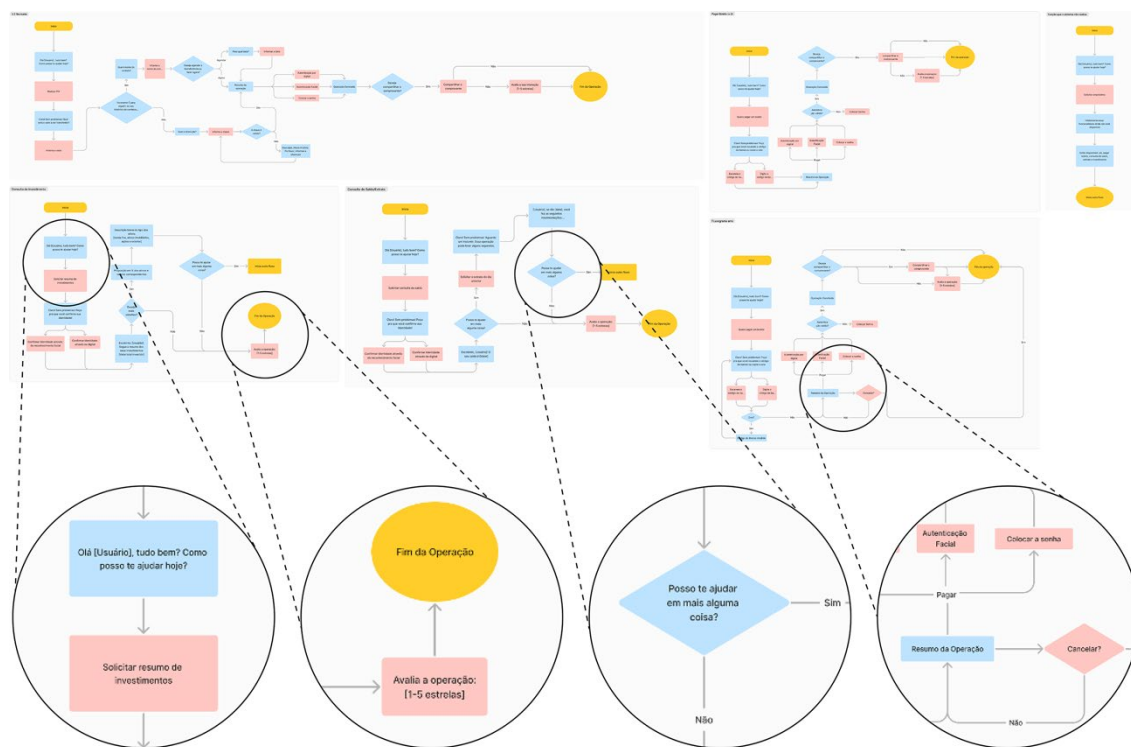
Com os requisitos definidos, iniciou-se o plano da Estrutura, onde buscou-se planejar o fluxo de navegação e as interações possíveis entre o usuário e a interface de voz, priorizando um design que fosse fácil e intuitivo para uso. Para tal, primeiramente desenvolveu-se uma estrutura conceitual que definiu a arquitetura da informação, organizando e categorizando o conteúdo do sistema de acordo com as funcionalidades obrigatórias (Figura 4).

Figura 4: Diagrama da arquitetura da informação. Fonte: autores (2025).



Então foram desenvolvidos seis fluxogramas com poucas opções e caminhos, que representavam as principais tarefas dos usuários: pagamento Pix, pagamento boleto, consulta saldo/extrato e consulta investimento, além de possíveis fluxos de erro ou indisponibilidade do sistema. Para facilitar a distinção entre o usuário e o sistema, foram utilizadas três cores diferentes: amarelo (início e o fim do fluxo), vermelho (*inputs*) e azul (*outputs* e *prompts*). Esses fluxogramas, exemplificados na Figura 5, foram criados para ajudar na visualização do passo a passo da interação, simulando uma conversa entre o usuário e a interface.

Figura 5: Fluxogramas com ampliação de detalhes. Fonte: autores (2025).

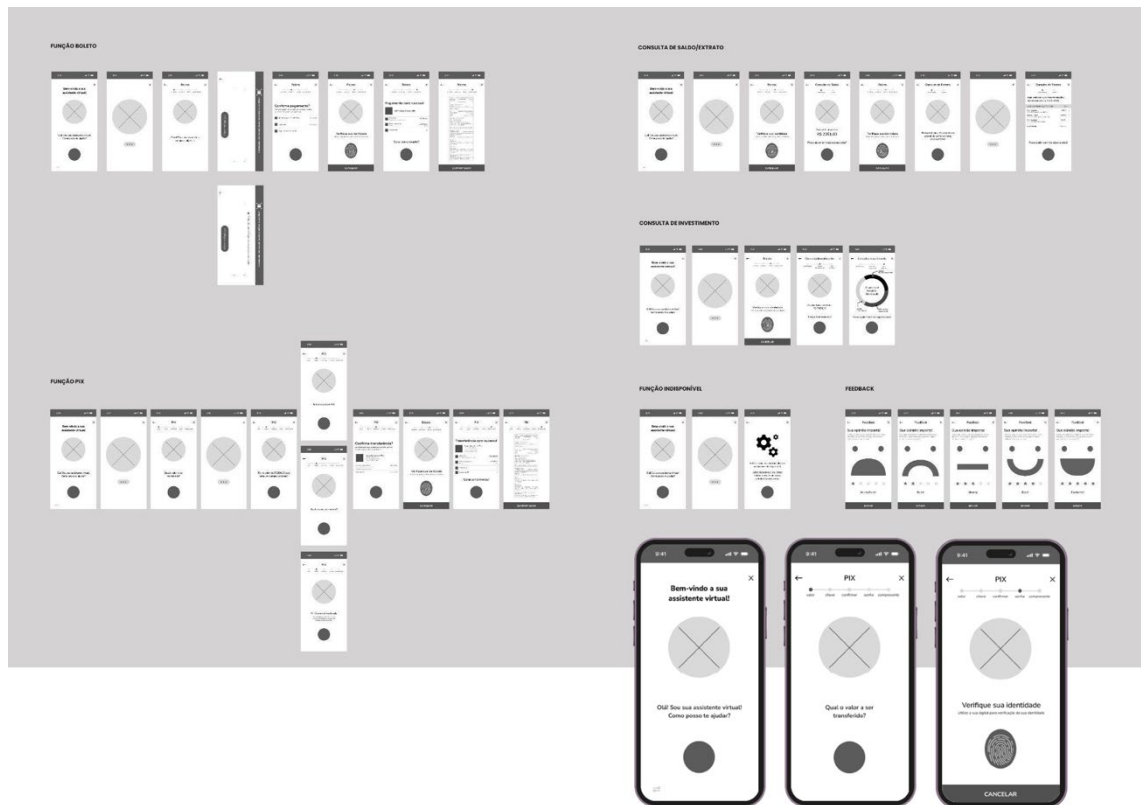


4.4 Esqueleto

Com o intuito de permitir uma interação multimodal, o foco do Esqueleto foi o desenvolvimento de *wireframes* e protótipos da interface gráfica que apoiaria a interação por comando de voz. Inicialmente, para assegurar que o produto seguisse convenções familiares, foi feita uma análise de referências, onde se pode observar a presença de um elemento gráfico que normalmente possui algum movimento, remetendo a uma forma orgânica, para representar a inteligência artificial.

A partir do que foi organizado e proposto no plano anterior, iniciou-se o rascunho das telas. Feitos os primeiros *wireframes* de baixa fidelidade, partiu-se para a confecção da média fidelidade no software de prototipagem. Nesta etapa foram criadas 27 telas em um protótipo digital, contemplando os fluxos apresentados anteriormente. Em síntese, buscou-se criar uma interface simples, que oferecesse um fluxo enxuto e focado na facilidade de uso, ajudando o indivíduo a evitar erros no processo (Figura 6).

Figura 6: Wireframes de média fidelidade. Fonte: autores (2025).



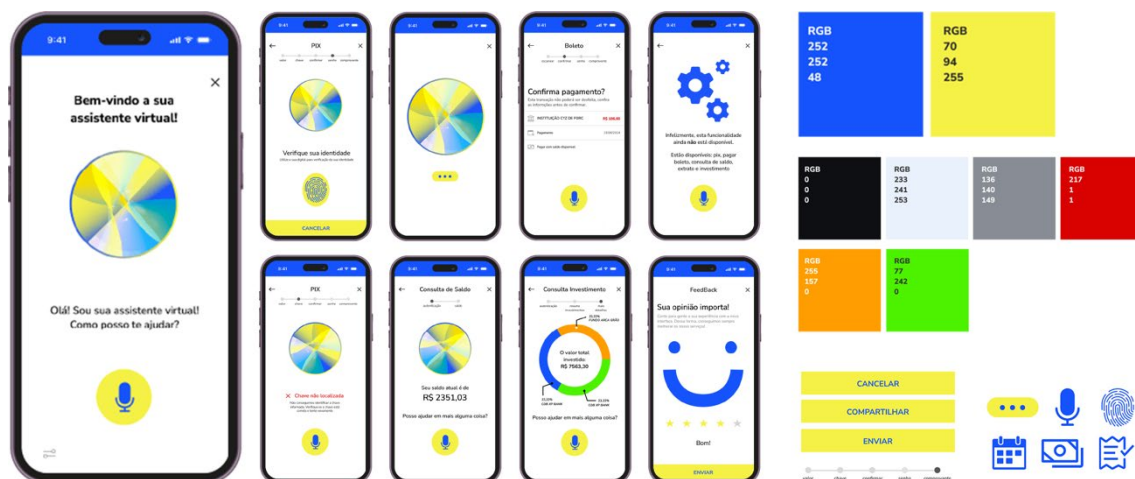
4.5 Superfície

O plano da Superfície teve como principal objetivo manter uma coerência visual com o atual aplicativo da instituição, assegurando uma integração com a interface gráfica do usuário. Para isso, foram utilizadas como referência o aplicativo “BB: Banco, Conta, Pix, Crédito” e o Manual de Identidade Visual do Banco do Brasil, essenciais para garantir que elementos familiares, como tipografia, cores e ícones, fossem preservados. Isso reforça a identidade visual e reduz a chance de estranhamento por parte dos usuários ao navegarem pela nova interface que, em sua maior parte, é utilizada através de comandos de voz.

A criação do elemento gráfico para identificar a assistente foi pensada para transmitir uma aparência dinâmica que mexe conforme o sistema fala, processa e ouve, remetendo a uma forma orgânica, prática recorrente nas interfaces por voz. Além disso, para orientar o usuário em cada etapa da interação, incluiu-se uma barra de progresso, facilitando a navegação e permitindo que a pessoa visualize facilmente em qual ponto do processo está.

O resultado global obtido foi um protótipo de alta fidelidade de uma interface conversacional para o Banco do Brasil, apoiada por uma expansão na sua interface gráfica, seguindo as convenções já utilizadas pela instituição. De modo geral todas as solicitações feitas para e pelo sistema acontecem por meio de comandos de voz, cujas respostas também são faladas. Nesse sentido, destaca-se que as informações apresentadas visualmente nas telas são um suporte, uma vez que são uma repetição das informações concedidas pelo sistema através da voz. A Figura 7 apresenta um compilado do resultado.

Figura 7: Principais telas da interface proposta. Fonte: autores (2025).



5 Discussões

A proposta desenvolvida nesta pesquisa foi orientada pelos cinco planos propostos por Garrett (2011), que permitem vislumbrar uma organização para o processo de design centrado na experiência do usuário. A adoção desse modelo, muito difundido no contexto do projeto de interfaces gráficas, se mostrou eficaz também no design de uma interface multimodal.

Nessa perspectiva, destaca-se que o plano da Estrutura demonstrou ser central para organizar os fluxos de interação de uma VUI. Os fluxogramas desenvolvidos buscaram respeitar a linearidade temporal e a lógica sequencial da interação por voz, propondo conversações diretas e enxutas, como recomenda a literatura (Dasgupta, 2018). A distinção cromática entre usuário e sistema contribuiu para a clareza na organização das interações. Sob a ótica do design da informação, essa etapa se destacou por prever cenários de *inputs* verbais e retornos do sistema, preocupando-se em uma compreensão gradual e controlada da informação pelo usuário.

Já o plano do Esqueleto possibilitou desenvolver a natureza multimodal da interface, que combinou interação por voz com elementos visuais, em consonância com a literatura que sugere o uso de múltiplos canais para reduzir a carga cognitiva e facilitar a compreensão das informações (Dasgupta, 2018; Cohen, Giangola & Balogh, 2004). Nessa dimensão, nota-se o papel das informações visuais na antecipação dos próximos passos pelo usuário e na confirmação dos *outputs*, o que pode minimizar dúvidas durante o uso. Consequentemente, a interface gráfica funciona como uma camada de apoio à interação por voz, contribuindo para a segurança do usuário ao confirmar ações e etapas.

Logo, o modelo de Garrett mostrou-se adequado não apenas para organizar o desenvolvimento da interface, mas também para lidar com os desafios específicos de um projeto multimodal, permitindo que a VUI se apoiasse em recursos gráficos sem comprometer sua autonomia. Isso evidencia como *frameworks* já consolidados podem ser aplicados com

eficácia no desenvolvimento de soluções contemporâneas que exigem articulação entre múltiplas formas de interação.

Por fim, a partir deste experimento projetual foi possível identificar boas práticas para o desenvolvimento de uma interface multimodal com foco na interação por voz, aplicada ao contexto bancário, corroborando os achados da literatura e algumas práticas de mercado. Assim, é importante:

- Manter a interação simples, evitando processos muito complexos. Sugere-se fragmentar os fluxos interativos em mais etapas, permitindo ao usuário se concentrar em um passo de cada vez, o que também simplifica a navegação;
- Durante o projeto, confeccionar fluxogramas consistentes e detalhados, utilizando uma linguagem gráfica padronizada, a fim de prever as principais possíveis interações entre o sistema e o usuário;
- Garantir que o sistema confirme as informações fornecidas pelo usuário para reduzir a carga cognitiva e permitir correções durante o processo;
- Permitir a personalização da interface para contemplar usuários com diferentes necessidades e níveis de experiência com este tipo de interação;
- Utilizar forma abstrata ou figurativa para representar o sistema na interação com o usuário, uma vez que grande parte dos sistemas conversacionais disponíveis no mercado apresenta este artifício e a literatura destaca a importância da personalidade na interface por voz;
- Por razões de privacidade e segurança, implementar sistemas que confirmem as operações bancárias solicitadas pelo usuário que não utilizem a fala, seja por meio dos digitais, escaneamento facial ou senha digitada, aliados a outros mecanismos como a localização do dispositivo.

6 Considerações finais

Esta pesquisa teve como objetivo propor uma forma de interação natural com o aplicativo do Banco do Brasil por meio de uma interface multimodal baseada em comandos de voz. Ao longo do desenvolvimento, foram aplicados conceitos de arquitetura da informação, design de interação, design da informação e design de interfaces, destacando o papel dos sistemas conversacionais na experiência do usuário. Os resultados mostram que o uso de assistentes de voz pode tornar as operações bancárias mais acessíveis e intuitivas, além de contribuir para a inovação no atendimento ao cliente. Dessa forma, este estudo reforça o potencial das interfaces de voz para atender às demandas de praticidade e inclusão em serviços bancários digitais.

Quanto às limitações da pesquisa, é preciso pontuar que a avaliação da interface foi restrita a uma avaliação heurística entre os planos do Esqueleto e da Superfície. Por isso, ainda que o protótipo de alta fidelidade apresente exemplos de respostas sonoras do sistema a comandos

simulados do usuário, recomenda-se, para futuras pesquisas, a realização de testes de usabilidade em um protótipo funcional, a fim de avaliar os resultados com maior profundidade, especialmente na validação da proposta em um contexto real de interação por voz. Outra fragilidade importante de sinalizar foi uso exclusivo de fontes secundárias a respeito do comportamento dos potenciais usuários. Mesmo que estes dados evidenciem padrões de mercado e contribuam para a construção de uma solução centrada em demandas reais, a ausência de fontes primárias de pesquisa com o usuário pode limitar a replicabilidade do estudo e não permite confirmar a validade dos resultados sem a realização de testes em futuras oportunidades.

Finalmente, apesar das limitações, entende-se que foi possível trazer contribuições pertinentes em um cenário de popularização das interfaces naturais. Em suma, este estudo apresentou uma proposta de interface multimodal para o aplicativo do Banco do Brasil, que recebe comandos por voz, responde por voz e exibe informações na tela como apoio, seguindo boas práticas para VUIs. Seus objetivos foram simplificar processos e otimizar operações financeiras, promovendo maior inclusão por meio de uma interface natural que utiliza a fala como principal meio de interação. Portanto, a partir deste experimento, espera-se também fomentar discussões a respeito do papel do design na comunicação efetiva da informação em interfaces por voz.

Referências

- Baer, K. (2021). *Information design workbook, revised and updated: Graphic approaches, solutions, and inspiration+ 30 case studies*. Quarry Books Editions.
- Banco Bradesco SA. (2024a). *Banco Bradesco*. [S. l.], 2024. Disponível em: <https://apps.apple.com/br/app/banco-bradesco/id336954985>. Acesso em: 30 out. 2024.
- Banco Bradesco SA. (2024b). *BIA Bradesco Inteligência Artificial*. [S. l.], 2024. Disponível em: <https://banco.bradesco/bia/>. Acesso em: 30 out. 2024.
- Banco do Brasil SA. (2024). *Banco do Brasil: Conta Digital*. [S. l.], 2024. Disponível em: https://play.google.com/store/apps/details?id=br.com.bb.android&hl=pt_BR. Acesso em: 30 out. 2024.
- Banco Inter SA. (2024). *Inter: Conta, Cartão e Pix*. [S. l.], 2024. Disponível em: <https://apps.apple.com/br/app/inter-conta-cartão-e-pix/id839711154>. Acesso em: 4 out. 2024.
- Banco Santander (Brasil) S.A. (2024). *Banco Santander Brasil*. [S. l.], 2024. Disponível em: <https://apps.apple.com/br/app/banco-santander-brasil/id613365711>. Acesso em: 30 out. 2024.
- Bouazid, A., & Ma, W. (2022). *The Elements of Voice First Style*. O'Reilly Media, Inc.
- Caixa Econômica Federal. (2024). *CAIXA*. [S. l.], 2024. Disponível em: https://play.google.com/store/apps/details?id=br.com.gabba.Caixa&hl=pt_BR. Acesso em: 30 out. 2024.
- Cohen, M. H., Giangola J. P., & Balogh, J. (2004). *Voice user interface design*. Addison-Wesley

Professional.

- Dasgupta, R. (2018). *Voice User Interface Design: Moving from GUI to Mixed Modal Interaction*. Hyderabad, Telangana, India. https://doi.org/10.1007/978-1-4842-4125-7_1
- Deloitte Global. (2023a). *Pesquisa Febraban de Tecnologia Bancária 2023 - Volume 1*. São Paulo: Febraban.
- Deloitte Global. (2023b) *Pesquisa Febraban de Tecnologia Bancária 2023 - Volume 2*. São Paulo: Febraban.
- Deloitte Global. (2024a). *Pesquisa Febraban de Tecnologia Bancária 2024 - Volume 1*. São Paulo: Febraban.
- Deloitte Global. (2024b) *Pesquisa Febraban de Tecnologia Bancária 2024 - Volume 2*. São Paulo: Febraban.
- Doherty, D.; & Curran, K. (2019). Chatbots for online banking services. *Web Intelligence*, v. 17, n. 4, p. 327–342.
- Garrett, Jesse James. (2011). *The elements of user experience: user-centered design for the Web and beyond*. 2. ed.
- Iliev, Y.; Ilieva, G. (2023). A Framework for Smart Home System with Voice Control Using NLP Methods. *Electronics* 2023, 12, 116. <https://doi.org/10.3390/electronics12010116>
- Itaú Unibanco Holding S.A. (2024). *Assistente Virtual Itaú*. Disponível em: <https://www.itaú.com.br/atendimento-itaú/para-voce/assistente-virtual-itaú>. Acesso em: 13 out. 2024.
- Klein, L. (2015). *Design for Voice Interfaces*. Sebastopol: O'Reilly Media, Inc.
- Neto, A. T., Fortes, R. P. M., & Silva Filho, A. G. da (2008). Multimodal interfaces design issues: the fusion of well-designed voice and graphical user interfaces. In *Proceedings of the 26th annual ACM international conference on Design of communication*, p. 277-278.
- Nu. (2024). *Nubank: conta, cartão e mais*. [S. l.], 2024. Disponível em: https://play.google.com/store/search?q=nubank&c=apps&hl=pt_BR. Acesso em: 30 out. 2024.
- Pearl, C. (2016). *Designing voice user interfaces: Principles of conversational experiences*. O'Reilly Media, Inc.
- Santos, A. dos. (2018). *Seleção do método de pesquisa: guia para pós-graduação em design e áreas afins*. Curitiba, PR: Insight.
- Sharp, H.; Rogers, Y.; & Preece, J. (2019). *Interaction Design: beyond human-computer Interaction*. 5. ed. Indianapolis: Wiley.
- Tomlin, W.C. (2018). What's a Persona?. In: *UX Optimization*. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-3867-7_2.
- Twyman, M. (1982). The graphic presentation of language. *Information Design Journal*, v. 3, n. 1, p. 2-22.

Sobre os autores

Maurício Elias Klafke Dick, Dr., FAAC/UNESP, Brasil <mauricioedick@gmail.com>

Anderson Massao Tanaka, Graduado, FAAC/UNESP, Brasil <tanaka.dsgn@gmail.com>